1-1-1997

The molecular and biological characterization of eight porcine respiratory coronavirus isolates

Dana Marie McCullough
Iowa State University

The molecular and biological characterization of eight porcine respiratory coronavirus isolates

by

Dana Marie McCullough

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Microbiology

Major Professors: Prem S. Paul and Ronald W. Griffith

Iowa State University

Ames, Iowa

1997

Graduate College
Iowa State University

This is to certify that the Master's thesis of

Dana Marie McCullough

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

This thesis is dedicated to my husband Russel

TABLE OF CONTENTS

INTRODUCTION

Porcine respiratory coronavirus (PRCV) is a member of the family *Coronaviridae*, and is closely related to transmissible gastroenteritis virus (TGEV) of swine. Coronaviruses are enveloped and have a positive-sense single-stranded RNA genome (Holmes, 1990). Transmissible gastroenteritis virus produces severe diarrhea in swine by infecting the villous enterocytes of the small intestine and causing them to slough off. This leads to malabsorption and dehydration. The mortality among piglets under 3 weeks of age has been shown to be 100% (Holmes, 1990; Saif and Bohl, 1986).

PRCV is believed to be a variant of TGEV because they are antigenically related. PRCV was first isolated in Belgium in 1986 (Pensaert et al., 1986) from a herd of pigs which tested positive for TGEV antibodies, but had never shown clinical signs of the disease. PRCV replicates in the respiratory tissue of swine with little clinical disease produced due to the virus itself, but it may predispose pigs to secondary infections (Halbur et al., 1993; Laude et al., 1993; Pensaert et al., 1986).

PRCV is now widespread in swine herds in Europe. It is interesting to note that with the emergence of PRCV in that region, the incidence of transmissible gastroenteritis has decreased (Pensaert et al., 1993). This may be because TGEV and PRCV are closely related and there is some cross-protection between the two viruses. The prevalence of PRCV in the United States is not known, but several researchers have isolated it in this country (Hill et al. 1989; Paul et al., 1992; Vaughn et al., 1995). TGEV and PRCV are genetically similar, and yet show differences in tissue tropism and pathogenicity. These characteristics make them good candidates for the study of coronavirus genes which are involved in tissue tropism and virulence (Vaughn et al., 1995).

There are many similarities between TGEV and PRCV. They both have the same genomic organization consisting of a polymerase gene, spike gene, ORF's 3, 3-1, 4, a matrix gene, a nucleocapsid gene and ORF 7 reading from the 5' to 3' end. PRCV has a deletion in the 5' end of the S gene when compared to TGEV which varies from 672 to 711 nucleotides depending upon the isolate (Britton et al., 1991; Laude et al., 1993; Rasschaert et al., 1990; Vaughn et al., 1995). The deletion in the S gene of PRCV is thought to play a role in the tissue tropism of the virus (Sanchez et al., 1992). There is also variation from TGEV in ORF-3 and 3-1 in PRCV. The PRCV isolates all have the consensus leader sequence preceding ORF-3, CTAAAC, altered or partially deleted which causes the subgenomic mRNA 3 to not be made in PRCV infected cells (Laude et al., 1993; Page et al., 1991; Wesley et al., 1990b). The 3 and 3-1 genes have been hypothesized to be involved in the virulence and pathogenicity of TGEV (Vaughn et al., 1995; Wesley et al., 1991). The PRCV isolates sequenced so far have alterations in these regions and cause little clinical disease in swine.

In our study we have obtained several isolates of PRCV and have characterized them to gain a better understanding of how they are related to TGEV, to examine the S, 3, and 3-1 gene deletions and alterations, and to determine if there are any differences between the isolates *in vitro* that may have in *vivo* significance. It has already been documented that isolates of PRCV from the United States have deletions in the S gene of varying sizes (Vaughn et al., 1995). In our study we have examined the S gene of four new isolates and mapped their deletions. The pathogenicity of the previous isolates has been explored (Halbur et al., 1994) and it was found that they vary in pathogenicity. The 3, and 3-1 genes of these isolates have also been analyzed to determine if there was a correlation between an intact gene 3 and virus pathogenicity (Vaughn et al., 1995). They found that the two PRCV strains that were the most

virulent had an intact gene 3 which was identical to the gene 3 of the virulent Miller strain of TGEV (Vaughn et al., 1995). This indicates that gene 3 may be linked to virulence, with an intact gene 3 causing stronger virulence.

We have analyzed the 3 and 3-1 genes of four new isolates to further confirm our previous observations. The biologic characteristics of these four new isolates, as well as four isolates that were used in previous studies, including plaque sizes, growth in two cell lines, and differences in the cytopathic effect of these viruses in the various cells were also examined to determine if there was potential correlation between genetic changes and *in vitro* characteristics.

LITERATURE REVIEW

Coronavirus History

The *Coronaviridae* family of viruses are large, enveloped, (+) strand RNA viruses containing only the genus Coronavirus. They have the largest genomes of all RNA viruses. Different species of coronaviruses have been implicated in many prevalent diseases of humans and animals, mainly causing respiratory or enteric illnesses (Holmes, 1990).

Avian infectious bronchitis was the first disease described that was caused by a coronavirus. It was first isolated in the laboratory by Beaudette and Hudson in 1937 (Beaudette and Hudson, 1937). The next coronavirus associated disease discovered was murine hepatitis virus in 1949 by Cheever et al (Cheever et al., 1949). In 1965, the first coronavirus was isolated from a human. It was isolated from a young boy with a cold by Tyrell and Bynoe (1965). Upon examination by electron microscopy they found that the virus they isolated from the boy was morphologically similar to avian infectious bronchitis virus (IBV). The virus was named the B814 strain (Tyrell and Bynoe, 1965). Around the same time, another group of researchers, Hamre and Procknow isolated five other virus strains from humans and found them to be morphologically identical to B814 and IBV as well (Hamre and Procknow, 1966). McIntosh et al. (1967) recovered six other strains of coronavirus from humans and showed the antigenic and morphologic relationships of the human viruses to murine hepatitis virus (MHV).

At this time, a new genus called Coronavirus was formed for these viruses based on their morphology. When viewed under an electron microscope, these viruses have a "corona", or crown-like appearance due to club-like projections of the surface of the virion (Tyrrel et al.,1968). This was the main criterion for this classification for some time. It is now possible to

characterize these viruses by other properties such as; structural proteins, number and size of mRNAs, antigenic cross-reactivity, sequence analysis, and homology with other known coronaviruses. Several coronaviruses have been isolated from a wide variety of animals such as rats, chickens, turkeys, pigs, dogs, cats and cattle (reviewed by McIntosh, 1990). It has recently been hypothesized that coronaviruses may potentially be important emerging pathogens due to their high rate of recombination, current species diversity, and unique replication strategy (Baric et al., 1995).

Coronaviruses

Coronaviruses are the only genus of the family *Coronaviridae*. They are large, enveloped, positive (+) strand RNA viruses. The virions are generally round, but are moderately pleomorphic. They measure 80-160 nm in diameter and are covered with club shaped projections which are approximately 20 nm in length (Holmes, 1990). These club shaped projections, called peplomers are composed of the S protein and give the virions their distinctive appearance. Classification of these viruses was at one time based solely upon the appearance of these "spikes" which are visible when negatively stained specimens are viewed under an electron microscope, but now they can be identified by various other characteristics (Holmes, 1990). Several viruses have been extensively characterized at the molecular level. These include mouse hepatitis virus, avian infectious bronchitis virus, transmissible gastroenteritis virus, bovine coronavirus, human respiratory coronavirus, and feline infectious peritonitis virus(Lai, 1990).

Coronaviruses exhibit strong species and tissue specificity, replication often being limited to epithelial cells of the respiratory and enteric tracts. Some replication has also been shown to occur in macrophages. The genome is 27-30 kb in size, which is the largest genome of all RNA viruses (Boursnell et al.,

1987). The RNA is non-segmented, capped and polyadenylated. Like other RNA viruses, naked genomic RNA is infectious when introduced into susceptible cells. Sequence analysis reveals the presence of at least 10 ORFs in the genome, some of which correspond to specific mRNAs (Lai, 1990). Based on the number of mRNAs, coronaviruses have 6-8 genes which encode for various structural and non-structural proteins. Although all of the mRNAs are polycistronic in structure, they are monocistronic in function.

There are several interesting features in the coronavirus genome. First, the gene order within the genome is comparable among different coronaviruses. MHV, BCV, and several other coronaviruses contain 2 additional genes which encode the hemagglutinin (HE) protein and a non-structural protein, p30. Neither of these genes are found in TGEV or IBV (Lai, 1990).

In general, the coding region of each mRNA contains one ORF, but there are some exceptions. The mRNA 1 of IBV and MHV has two slightly overlapping ORFs in the 5' end. They are translated into a large protein by a mechanism of ribosomal frameshifting. This mechanism requires a pseudo-knot structure around the overlapping region. Gene 1, which encodes for the RNA dependent RNA polymerase accounts for more than two-thirds of the genome (Lai, 1990). In the 5' end of the genome there are 60-70 nucleotides termed the leader sequence. A portion of the same sequence is found at the 5' end of every mRNA. This sequence is believed to be important in the regulation and transcription of genomic RNA and sub-genomic mRNAs. At every intergenic region there is a small stretch of a consensus sequence of 8-10 nucleotides which is important for the transcription of downstream genes (Lai, 1990). The molar amounts of each mRNA produced are different. The smaller mRNAs generally are more abundant than the larger mRNAs.

There are 3 or 4 structural proteins which make up the coronavirus virion; the N, or nucleocapsid protein, the M, or matrix protein, the S, or spike protein

and some coronaviruses contain the HE protein on the external membrane surface. Molecules of the N protein surround the genomic RNA and form a long, flexible nucleocapsid 6-8 nm in diameter, with a helical shape (Holmes, 1990). The nucleocapsid is surrounded by an envelope which is derived from the Golgi apparatus or the rough endoplasmic reticulum (RER) (Holmes, 1990). The envelope has a typical lipid bilayer structure, which contains the two other viral proteins, M and S.

Another viral protein, hemagglutinin, or HE, is found in the envelope of some coronaviruses (Callebaut et al, 1980). It binds to the neuraminic acid residues on cell membranes and causes hemagglutination (Holmes, 1990).

The M protein is a transmembrane glycoprotein, 20-30 kDa in size which differs from the glycoproteins in other virus groups. It is deeply embedded in the envelope so only a small region is exposed to the outer membrane surface. The amino acid sequence of this protein suggests that it crosses the lipid bilayer three times and has a large domain which lies beneath the bilayer (Holmes, 1990). It may bind the nucleocapsid to the viral envelope during assembly and this may be why the virion buds from the Golgi instead of the plasma membrane (Holmes, 1990). Antibodies to the M protein can neutralize the virus only with the help of complement.

The S glycoprotein is 180-200 kDa and is the structural protein of the spikes on the viral envelope. It has a small anchor in the lipid bilayer, but most of the molecule lies outside of it. It is transported to the plasma membrane of infected cells. A comparison of the nucleotide sequence of the S genes of several coronaviruses reveals some interesting features. There is a short cytoplasmic domain at the carboxy terminus of the S gene which is rich in cystine residues. This region may have a complex tertiary structure which could play a role in the assembly of the virions. Considerable diversity has been found in both the length and nucleotide sequence of the amino-terminal ends of the S

proteins in different viruses. In some viruses, such as MHV, this protein is cleaved into two 90 kDa proteins. This cleavage is necessary for the infectivity of these viruses (Sturman, 1980). These differences, along with insertions in the S gene coding region may result from the high frequency of recombination during replication among these viruses. The biological functions of the S protein are diverse. It is responsible for binding to the virus receptor on target cells, and has been shown to induce cell fusion (Sturman et al., 1983). Antibodies to it can neutralize viral infectivity and it is believed to play an important role in coronavirus pathogenesis.

The N protein is an internal component of the virion. It is a phosphoprotein of 50 kDa which binds to the genomic RNA and provides a structural basis for the helical nucleocapsid of the virion (Lai, 1990). It binds to a specific segment of the leader RNA and may provide not only a structural function, but also a regulatory function for RNA synthesis (Stohlman et al., 1988).

Another coronavirus glycoprotein, HE is a hemagglutinin which has been found in BCV, HEV, HCV-OC43, TCV and some MHV strains, but has not been found among viruses in the antigenic groups 1 and 3 (Holmes, 1990). It is believed to make up the smaller spikes seen on some viruses under electron microscopy (Dea et al, 1988).

There is some similarity in function between the HE protein and the HA protein of the influenza C virus. Monoclonal antibodies to the HE protein can inhibit virus induced hemagglutination and neutralize virus infectivity (Deregt et al., 1987).

There are four distinct antigenic groups of coronaviruses, as shown in Table 1 (Holmes, 1990). Viruses within groups show partial antigenic cross-reactivity, but they are distinguishable by host specificity and clinical syndromes. For example, serum from an animal exposed to feline enteric coronavirus

Table 1. Coronaviruses: Antigenic Groups[a]

Antigenic Group	Virus	Host
I	HCV-229E	Human
	TGEV	Pig
	CCV	Dog
	FECV, FIPV	Cat
II	MHV	Mouse
	HEV	Pig
	BCV	Cow
III	IBV	Chicken
IV	TCV	Turkey
New group	PEDV	Pig

[a] Adapted from Holmes (1990).

(FECV) cross-reacts with TGEV, CCV, and FIPV (Ingersoll and Wylie, 1988). This cross-reactivity does not confer protective immunity, however.

There are large numbers of serotypes among most coronaviruses due to the high rate of recombination and the general frequency of mutations among RNA viruses. Recombination, which is one cause of the high mutation rate, occurs in approximately 10% of the progeny virions (Baric et al., 1995). There is also a rather high frequency of mutation among the virions due to the lack of proofreading ability in the RNA dependent RNA polymerase. This natural selection of virus variants in vivo may account for the large number of serotypes.

Coronaviruses can cause either persistent or cytocidal infections of cells. The virus strain and the host-cell type determine the type of infection. Cytocidal infections cause the cells to either form multi-nucleated syncytia or cause direct cell lysis (Holmes, 1990). Persistent infections have been observed in cell

culture, where the cells may produce virus for several weeks with no CPE or cell death (Holmes, 1990). Most viruses show a marked tissue tropism and only grow in cells from the host species, however in some studies the species border has been crossed with highly passaged strains of virus (Baric et al., 1995).

Coronavirus replication takes place in the cytoplasm of infected cells. The virions attach to specific receptors on the plasma membrane of the host cell by means of the S protein (Boyle et al., 1987). In some viruses, HE also plays a role in this attachment. It is not yet clear how coronaviruses penetrate the cell membranes, but it is believed to involve the endosomes. Virus entry by cell fusion has also not been ruled out.

The first event after penetration is the uncoating and attachment of the genomic RNA to the ribosomes where virus specific RNA dependent RNA polymerase is made. This is not normally present in the cell and must be synthesized by incoming viral genomic RNA. Inhibition of protein synthesis at any point in the replication cycle will interrupt viral RNA synthesis. The (+) strand genomic RNA is transcribed to make complementary, full length, minus (-) strand RNA which has a poly(U) sequence at the 5' end. This negative strand serves as a template for new genomic RNA and also 5-7 subgenomic mRNAs. These subgenomic mRNAs form a nested set of overlapping molecules with common 3' ends. Only the 5' sequence of the mRNA, that is not found in the next smaller mRNA, is translated. At the 5' end of the genomic RNA there is a leader sequence of about 60-70 bases. This same sequence is also found at the 5' end of the subgenomic mRNAs.

There have been three models proposed for the transcription of the mRNAs. One model suggests a separate initiation site for each mRNA, but a common termination site. Another model suggests that the mRNAs are derived from the cleavage and splicing of a precursor full-length RNA product. It has also been hypothesized that the mRNAs are synthesized using subgenomic (-)

strand template RNA, all of which have the same 5' end but terminate at separate points. All of the proposed models utilize a discontinuous transcriptional process. This could occur through several mechanisms. The first model involves a "looping out" of the template RNA to allow the RNA polymerases to jump from the leader region to an internal initiation site. The second model involves post-transcriptional processing, a method in which the leader is fused to the mRNAs by a splicing mechanism. This is thought to be highly unlikely, as there are no consensus splicing sequences at the leader and mRNA junction sites. The leader primed transcription theory is thought to be the most likely. The leader mRNA is transcribed at one end of the template, disassociates, and then rejoins the template RNA downstream to serve as a primer for transcription. For all of the above models, template RNA is of genomic length, but subgenomic lengths of RNA have been found. A majority of the evidence supports the leader primed theory of transcription. Several small leader sequence RNAs have been detected in the cytoplasm of MHV infected cells.

Only the 5' end of mRNA is translated, thus each mRNA yields only a single polypeptide. The gene order of arrangement in general is 5'-pol-(HE)-S-M-N-3'. There are several additional open reading frames for small non-structural proteins located in various places along the genome for different viruses. The biological significance of these ORFs and their products is not yet understood, but they may be linked to pathogenesis. The nucleocapsid protein and several of the non-structural proteins are synthesized on polysomes in the cytoplasm. The HE, S, and M proteins are synthesized on polysomes which are attached to the RER. The accumulation of viral structural proteins is a prerequisite for the assembly of virions. It may also signal the switch of the viral RNA from making mRNAs to genomic RNA synthesis for packaging into mature virions. The assembly of the S protein occurs in the RER, then it is transported

to the plasma membrane. The accumulation of these proteins on the membrane can make the cell susceptible to lysis by antibodies and complement. It is also required for cell fusion. The M protein is also synthesized on the ER. It has several hydrophobic domains which span the membrane three times. The insertion sequence can be either the first or the third transmembrane sequence. In TGEV, a signal peptide that has 17 amino acids is cleaved from the amino terminus of the M protein in the ER, then the M protein is transported to the golgi.

Virion assembly mainly occurs in the cytoplasm. The helical nucleocapsid is formed in the cytoplasm by an interaction of genomic RNA with the N protein. Assembly by budding initially occurs at membranes between the RER and golgi, via the cytoplasmic domain of the M protein. Complete virions are able to assemble in the cell before the S protein has accumulated at the plasma membrane and made the cell susceptible to attack. This may play an important role in coronavirus persistence. The virions are released from the cells in two ways. They can either cause the cells to lyse, releasing the virions or they can be released from intact cells by making use of cellular secretory pathways.

It has been suggested by Baric et al. (1995) that coronaviruses could be potentially important emerging pathogens. The large genome size and unique replication strategy coupled with the high rate of recombination during mixed infections gives the coronaviruses a high capacity to evolve (Baric et al., 1995). The majority of emerging RNA viruses are potential zoonotic pathogens, human immunodeficiency virus, hanta virus and influenza being prime examples. There is little known about the molecular mechanisms which mediate virus spread between species, but the high mutation rates in coronaviruses suggest a large quasispecies population. Baric et al. (1995) conducted research that examined the polymerase error rates and RNA recombination frequencies in MHV. They found that there was a mutation rate of 10^{3}-10^{5} per round of replication. The

most likely model to explain the high, progressively 5'-3' increased recombination rate in MHV is from the large genome size, the discontinuous method of transcription and the presence of transcriptionally active full and subgenomic length plus and minus strand RNAs which increase the amount of template for strand switching (Baric et al., 1995). Baric et al. (1995) were also able to show that MHV can cross the species barrier by getting it to grow in hamster baby kidney cells in relatively few passages. There are newly recognized animal coronaviruses which incorporate parts of other viral genomes such as porcine epidemic diarrhea virus, which has a portion of its sequence that is similar to that of human coronavirus 229E. This is another example of the coronaviruses as important emerging pathogens. The above evidence shows the ability of coronaviruses to evolve rapidly and possibly bridge species barriers.

Transmissible Gastroenteritis Virus

TGEV belongs to the *Coronaviridae* family of viruses. It was first isolated in 1946 in Europe by Doyle and Hutchings. The virus causes transmissible gastroenteritis which is a highly contagious enteric disease of swine characterized by severe diarrhea, vomiting and dehydration in susceptible pigs. It is often fatal in young pigs, with a mortality rate near 100% in pigs under 2 weeks old (Saif and Bohl, 1986). The disease is a problem in nurseries in the United States, with conservative estimates of annual economic losses to the pork industry between 25 to 75 million dollars (Miller et al., 1982).

There are currently no ideal methods for treatment, prevention or control of TGEV. Serological tests that identify antibodies to TGEV in the sera of infected animals have been problematic due to the emergence of PRCV. The antigenic cross-reaction among isolates of TGEV and PRCV has been discussed by many researchers (Callebaut et al., 1988; Garwes et al., 1988;

Wesley et al, 1990b; Vaughn and Paul, 1993). There is only one serotype of TGEV, but monoclonal antibodies have been used to detect variations among TGEV isolates (Laude et al. 1986, Vaughn and Paul, 1993; Zhu et al., 1990). This may be one of the reasons commercial vaccines have failed to protect some herds. Autogenous vaccines have been shown to be effective in some cases (Paul, 1988).

TGEV has several antigenic properties in common with other coronaviruses. There are two major antigenic groups of coronaviruses (Holmes, 1990). TGEV , feline infectious peritonitis virus (FIPV) and canine coronavirus (CCV) have been shown to be closely antigenically related and are in the same group (Pensaert et al. 1991). Cross-reactivity has been shown among these viruses by serological tests (Sanchez et al. 1989). PRCV is also in this group and is believed to be a respiratory variant of TGEV, and cannot be distinguished from TGEV accurately in serological tests. TGEV and PRCV are antigenically unrelated to another known porcine coronavirus, porcine epidemic diarrhea virus (PEDV).

The genome of TGEV is similar to other coronaviruses. It has a positive-sense RNA genome of 28.5 kb, which can be directly infectious when introduced to susceptible cells (Lai, 1990; Holmes, 1990). TGEV has three main structural proteins, the S, M, and N. The S, or spike protein forms the peplomers on the viral surface which are responsible for virus attachment. It has a molecular mass of approximately 200 kDa and is glycosylated (Garwes et al., 1976). It is believed to be the main protein involved in antigenicity and tissue tropism (Laude et al., 1993; Rassachert et al.,1990; Wesley et al., 1990a). Antibodies to this protein neutralize the virus and prevent viral infection (Garwes et al., 1978).

The M or matrix protein is found entwined in the viral envelope. It has a molecular mass of 25 to 30 kDa and is also glycosylated. Structure analysis shows that it passes through the membrane three times, with only a small

portion actually being exposed to the external surface. Antibodies to this protein do not neutralize virus infection. The N, or nucleocapsid protein complexes with the genomic RNA to form the nucleocapsid inside the viral envelope.

Two main types of TGEV infections are seen in swine herds, the epizootic and the enzootic forms. An epizootic infection occurs when the TGEV virus is introduced to a susceptible herd. There is a rapid spread of the virus to pigs of all ages. Many animals will lose their appetite, become lethargic, and develop vomiting and/or diarrhea (Bohl, 1989) Mortality can run as high as 100% in neonates, but the severity of the disease declines with age. The disease usually is self limiting, and terminates within a few weeks if no new susceptible pigs are brought into the herd (Bohl, 1989). The enzootic form of TGE refers to the persistence of the infection and disease within a herd. This usually occurs in herds where susceptible animals are continuously being brought in, as in a continuous farrowing operation (Bohl, 1989). The sows in these herds are commonly protected enough to not show severe clinical disease and also to provide some lactogenic immunity to their piglets, but at weaning, when maternal antibodies decline the piglets break out with the disease. Mortality at this stage is usually less than 20% due to the age of the piglets but weight gain suffers (Bohl, 1989).

TGEV infects the enteric tract of swine, replicating in the absorptive enterocytes (Sirinarumitr et al., 1996). TGEV binds to the cell through two receptors, aminopeptidase-N (CD13) (Delmas et al. 1990) and another 200 kDa receptor (Weingartl and Derbyshire, 1993). By using *in situ* hybridization it has been shown that there is TGEV nucleic acid present in the crypt epithelial cells as well, although these cells do not have either of these receptors. It has been speculated that TGEV may bind non-specifically to the cell membrane and then infect them (Weingartl and Derbyshire, 1993). It could also be that a TGEV infection in these cells is non-productive and that is why nucleic acid is seen, but

protein is not (Woods et al., 1981; Saif and Wesley, 1992). The infection of the villous enterocytes results in cell destruction and shedding, which causes villous atrophy within the small intestine (Bohl, 1989). This leads to the inability to absorb some nutrients and disturbs the chemical balance within the intestine causing diarrhea and digestive problems (Moon, 1978). The disease is not as fatal in older pigs in part for their ability to replenish the villous enterocytes rapidly, taking only 2-4 days in 3 week old pigs as opposed to 7-10 days in one day old pigs (Moon, 1978).

Because of the high mortality rate in young piglets, one of the goals of researchers has been to develop a vaccine that provides lactogenic immunity. It has been shown that exposure of the sow to virulent virus provides lactogenic immunity to the piglets (Bohl and Saif, 1975). This immunity is not carried post weaning. There are vaccines available and licensed for TGEV which consist of either inactivated or live attenuated virus. They seem to provide variable protection (Moxley and Olson, 1989).

A small plaque variant was developed which was derived from a persistently infected swine leukocyte cell line. This isolate was avirulent for 3 day old piglets and pregnant gilts. The small plaque mutant could not be isolated from pigs after inoculation, but the pigs developed virus neutralizing antibodies against TGEV in serum, colostrum, and milk. Only 62% of the piglets were sick after challenge with the virulent Miller strain and 14% died (Woods, 1978). This isolate was found to have a 462 nucleotide deletion in the 3 and 3-1 region of the genome, but had an intact S gene. Researchers hypothesized that the 3 and 3-1 regions are important in the virulence and pathogenesis of TGEV (Westly et al. 1990a).

Porcine Respiratory Coronavirus

PRCV is antigenically and genetically related to TGEV. It replicates in the respiratory tract of swine, and causes little to no respiratory disease. It is found all across Europe, and to a lesser extent the United States and Canada.

PRCV was first isolated in Belgium in 1984, when it was discovered that there were exceptionally high numbers of swine which were seroconverting to TGEV without evidence of clinical disease in the herd (Pensaert et al., 1986). A coronavirus isolate designated TLM 83 was obtained from the respiratory tract of swine in which no enteric infection occurred, but seroconversion did (Pensaert et al. 1986). The virus was later renamed porcine respiratory coronavirus due to its tropism for the respiratory tract. PRCV was first isolated in the United States in 1989 by Hill et al. (1989).

PRCV was initially isolated in primary pig kidney cells, but the CPE rapidly disappeared and the cells recovered. Other cell systems were utilized to optimize the growth of PRCV. It replicates well in porcine continuous cell lines ST and PD5, causing syncytia formation and destruction of the monolayer (Pensaert, 1989).

PRCV spread quickly across Europe and has recently been found in the United States and Canada. PRCV was enzootic in Belgium in 1986 with 100% of swine farms having been infected (Pensaert, 1989). It has been suggested that it may spread aerogenically due to the infiltration of the virus into areas which had no incidence of TGEV such as Denmark. Infection has been known to occur in herds which are very isolated from other herds. Also, infection is more common in the autumn months in Europe, where the foggy and rainy weather further facilitate aerosol spread (Laude et al. 1993). The possibility of other modes of transmission including fomite contamination and spread by stray animals are still under investigation.

The differentiation of TGEV and PRCV by traditional virus neutralization tests is not possible due to their common antigenicity. TGEV and PRCV are both neutralized by serum to each other. This has raised some concerns with the export of pigs from countries with PRCV to countries which require pigs to be TGE-free before being exported (Pensaert, 1989). Monoclonal antibodies have been developed which react to a non-neutralizing portion of the S gene which is present in TGEV, but not in PRCV. These monoclonal antibodies have been utilized in competitive ELISAs which differentiate between TGEV and PRCV antibodies in serum (Callebaut et al.; 1989; Garwes et al., 1988; Jabrane et al, 1992). The presence or absence of blocking antibodies determines whether the animal has seroconverted to TGEV or PRCV (Callebaut et al., 1989).

Other tests have been developed for the differentiation of TGEV and PRCV as well. Complementary DNA probes have been used to differentiate TGEV and PRCV infection in tissue samples (Wesley et al, 1991). Recently, a test has been developed utilizing a radiolabeled probe and an *in situ* hybridization for the differentiation of TGEV and PRCV in paraffin embedded formalin fixed samples and tissue culture (Sirintumatr et al., 1996). Further work is being done with this project to develop a similar test which uses a non-radiolabeled probe instead to make the test safer and more practical in the diagnostic lab. These tests, with the exception of isolation from nasal swabs in tissue culture, are all postmortem tests. Accurate serological tests need to be developed.

Piglets obtain maternal antibodies through colostrum shortly after birth, and the level of antibody rapidly declines. The amount of maternal protection to TGEV conferred to the piglets depends upon how long prior to birth the sow was infected with PRCV and the age of the piglet during challenge. Whatever the immunization procedure in experimentally infected pigs, no correlation was seen between the antibody level and the degree of passive protection transferred to

the piglets (Bernard et al., 1989). In herds with endemic infections of PRCV, the antibody titers in the young pigs increase during suckling, decrease upon weaning and increase again after 4-8 weeks of age. This indicates that the piglets must become infected with PRCV sometime after weaning (Pensaert, 1989). Natural infection with PRCV induces protective lactogenic immunity against TGEV, however protection is not complete. In a study by Bernard et al. only two of seven litters from sows infected with PRCV didn't show morbidity after challenge with TGEV. Five of seven litters had no mortality (Bernard et al., 1989). The age of the piglets at challenge may be a relevant factor as age related resistance to TGEV is well documented. Other researchers have found conflicting data showing that lactogenic immunity from sows infected with PRCV may not adequately protect their litters against a natural TGE challenge (Hooyberghs et al., 1991). This may be due to the conditions used for evaluation of protective immunity (Bernard et al., 1989). Cases of clinical TGE have decreased in Europe concurrent with the seroconversion to PRCV. This can be considered as an argument in favor of cross-protection (Bernard et al. 1989).

PRCV has been shown to replicate in the nasal mucosa, trachea, lungs, tonsils, bronchial and mesenteric lymph nodes and to a small extent in the small intestine. It replicates to high titers in the respiratory tract. Replication mainly occurs in alveolar cells, but also in epithelial cells of nasal mucosa, trachea, bronchi, bronchioli and in alveolar macrophages (Cox et al., 1993). After the primary replication cycle, viremia occurs. The virus is sometimes able to reach the intestinal tract, possibly due to the swallowing of virus from the respiratory tract or from the viremia (Cox et al., 1993). PRCV has been shown to have limited replication in the intestinal tract which is limited to a few cells (Cox et al. 1993; Halbur et al. 1993; Sirinarumitr et al., 1996). PRCV has experimentally

been isolated from the caudal small intestine when the virus was inoculated intravenously or directly into the intestinal tract (Cox et al., 1993).

Experimental infection of the respiratory tract do not always cause clinical signs, but histological evidence is present (Halbur et al., 1993). PRCV isolates in Europe caused a mild to moderate bronchiointersitial pneumonia in neonatal pigs, however the pigs were clinically normal (Cox et al., 1990a; O'Toole et al., 1989). American isolate AR-310 was shown to be pneumopathogenic for 3 day old gnotobiotic pigs. However, there was no clinical disease seen. It is thought that infection with PRCV may enhance disease signs when concurrently infected with other respiratory agents (Halbur et al., 1993; Pensaert, 1989). The lesions which are seen in the lungs may predispose the pigs to secondary bacterial infections (Halbur et al., 1993). Studies have been done which look at the pathogenicity of concurrent infections with PRCV and swine influenza virus (Lanza et al., 1992; VanReeth and Pensaert, 1994). These studies found concurrent infections with PRCV and two strains of swine influenza virus did not enhance disease and indicated that there may be possible interference between the two viruses (Lanza et al., 1992). Porcine reproductive and respiratory syndrome (PRRS) virus has also been isolated from herds with PRCV, and in many clinical cases of PRCV there was a mixture of bacteria isolated from the respiratory tract (Halbur et al., 1993).

The decrease in the incidence of TGEV in Europe concurrent with the spread of PRCV suggests that pigs infected with PRCV were immune to infection with TGEV. However, conflicting results have been obtained in different experimental studies by many investigators. Cox et al. (1993) showed that pigs infected with PRCV and challenged with TGEV show signs of diarrhea and TGEV was still shed in feces. They concluded that infection with PRCV does not induce intestinal immunity to TGEV, but does offer some protection (Cox et al., 1993). There was a secondary immune response to TGEV in most of

the PRCV immune pigs, but they all showed evidence of TGEV replication, although clinical signs were not as severe as those in the control pigs (Cox et al., 1993). It is posoble that a decrease in the incidence of TGEV is concurrent with changes in farming practices as well. Antigenic reactivity between TGEV and PRCV has been shown by immunoblots with monoclonal antibodies to the M, N and S proteins. Some of the non-neutralizing monoclonal antibodies to the S protein failed to react to PRCV, indicating a difference between the viruses in this region (Callebaut et al., 1988).

It is unknown how PRCV originated. It has been speculated that it is an attenuated TGE virus from a vaccine strain. Studies have shown that upon repeated passage in continuous cell lines TGEV can become adapted to replicate in pulmonary cell lines. Sequence analysis shows that TGEV and PRCV are identical in most regions, substantiating the more widely held belief that PRCV is a TGEV variant (Pensaert, 1989). The genome of PRCV is organized similar to TGEV. It has a polymerase gene, S, M and N genes. The mRNAs 2 and 3 are smaller in PRCV. Researchers in Europe sequenced these regions in two Belgian PRCV isolates and one French isolate and found that there was a 672 nucleotide deletion which occurred 59 nucleotides downstream of the S gene initiation site when compared to TGEV (Britton et al, 1991). Deletions in the S protein may be a way of altering the tropism and concurrently the pathogenicities of these viruses (Britton et al., 1991). The mRNA 3 was smaller due to the presence of a new RNA-leader binding site upstream of the PRCV ORF-3 gene. Four other ORFs were identified and shown to be 98% similar to TGEV ORF-4, M, N and ORF-7. No other deletions or specific PRCV sequences were identified (Britton et al., 1991). PRCV and TGEV genomes differ only by a few deletions and point mutations. A comparison of these sequences with TGEV showed that the leader RNA sequences involved in transcription were the same (Page et al., 1991). It also showed that mRNA 2

and 3 were smaller due to the reasons outlined above (Britton et al., 1991; Page et al, 1991). Studies done on the mRNA 3 identified several small deletions in the region which resulted in a loss of the potential TGEV ORF-3a gene (Page et al., 1991). The PRCV mRNA 3 is approximately 200 nucleotides smaller than the equivalent in TGEV. There are several small deletions, 84 nucleotides total in the PRCV genome which correspond to the 5' end of TGEV mRNA 3, two of these deletions being in the TGEV 3a gene region resulting in a loss of this potential gene in PRCV (Britton et al., 1991). The exact mechanism for the deletion events is not known. The TGEV ORF-3a gene identified in three different strains of TGEV is present as a pseudogene in PRCV. The small plaque mutant of TGEV had a 462 nucleotide deletion which eliminated the ORF-3a in this virus (Woods, 1978). These observations show that the protein encoded by ORF 3 is not needed for replication of TGEV, but its role in pathogenicity has yet to be determined (Britton et al., 1991). It is believed that the 3 and 3-1 genes play a role in the pathogenesis because it is altered in all isolates of PRCV studied so far. Deletions from the US isolates differ from the European isolates, but they show the same general pattern; the S gene is truncated and ORF-3 is converted to a pseudogene (Laude et al., 1993). Further characterization of the US isolates needs to be done to determine the possible roles that these mutations play in virulence.

THE MOLECULAR AND BIOLOGICAL CHARACTERISTICS OF EIGHT PORCINE RESPIRATORY CORONAVIRUS ISOLATES

A paper to be submitted to
Veterinary Microbiology
Dana M. McCullough and Prem S. Paul

Abstract

Four new isolates of porcine respiratory coronavirus (PRCV) were isolated from various swine herds in the United States. Biological characteristics of these viruses, as well as four other PRCV isolates, were examined. Studies included growth properties in two cell lines, plaque sizes, and cytopathic effects. Growth curves showed a similar pattern of growth among the isolates within the swine testis and porcine intestinal cell lines. There is an initial peak at around 24-30 hours post infection, then a slight decline with a second peak at approximately 48 hours post infection. Overall, the titer of viruses was higher in the intestinal cell line, however isolate IA1894 had a lower titer than the others. It was also the only isolate which was non-cytopathic in this cell line. Plaque sizes among the isolates varied. The PRCV isolate IA1894 had the smallest plaque size, the average being 1.0 mm, while PRCV isolate 725 had the largest plaque size, the average being 5.2 mm.

The nucleotide sequence of the 5' end of the S genes, the 3 genes 3, and 3-1 of the four new PRCV isolates were analyzed and compared to other PRCV and TGEV isolates. All four of the PRCV had a large deletion in the 5' region of the S gene. The size of the deletion varied from 621 to 711 base pairs. The PRCV isolate 725 had a 621-nucleotide deletion starting 47 nucleotides after the

S gene start site similar to that for isolate LEPP and AR310. The PRCV isolate PON had a 681-nucleotide deletion starting 62 nucleotides after the S gene start site, identical to the previously characterized PRCV isolate ISU-1. The PRCV isolate 306 had a 675-nucleotide deletion starting 54 nucleotides after the S gene start site, while the PRCV isolate NVSL 5170 had a 711-nucleotide deletion 27 nucleotides after the S gene start site. This is the largest deletion reported at this time.

Analysis of the nucleotide sequence of gene 3 of the four isolates showed some differences among the isolates. All of the isolates had a CTAAAC leader RNA binding site and the ATG start codon for the gene 3. The gene 3 of 725 and NVSL 5170 were predicted to yield a protein of 72 amino acids, which is the same size as PRCV isolates AR310 and LEPP and also the virulent Miller strain of TGEV. The PRCV isolate PON was predicted to yield a truncated protein of only 16 amino acids due to a change in sequence which coded for a stop site 48 nucleotides from the start site. The PRCV isolate 306 was markedly different from the other isolates in that it had a deletion of 7 nucleotides, which caused a frameshift and an insertion of 29 nucleotides which coded for a stop codon. The predicted protein size was 66 amino acids.

There was high variability in the gene 3-1 of all the isolates as well. The gene 3-1 of 725, NVSL 5170, and PON were all preceded by the CTAAAC leader binding sequence. The PRCV isolate 725 was predicted to yield a truncated (63 amino acids compared to 244 amino acids in the virulent Miller strain of TGEV) 3-1 protein due to a 1 nucleotide deletion. This is similar to that for isolates AR310 and LEPP. The PRCV isolate PON was predicted to yield an 80 amino acid protein due to a deletion of two nucleotides. Isolate 306 was totally different from the other isolates. It had no CTAAAC leader binding sequence and no ATG start site due to a 52-nucleotide deletion which leads to the prediction that there is no 3-1 protein formed. The information gained from

these new PRCV isolates should be useful in gaining a better understanding of the pathogenesis of PRCV.

Introduction

Porcine respiratory coronavirus (PRCV) and transmissible gastroenteritis virus (TGEV) are members of the *Coronaviridae* family of viruses (Saif and Bohl, 1986; Pensaert, 1989). Coronaviruses are enveloped and have a positive-sense, single-stranded RNA genome (Holmes, 1990). Transmissible gastroenteritis virus (TGEV) of swine causes severe diarrhea with high mortality in piglets under 3 weeks of age (Holmes, 1990; Saif and Bohl, 1986). It infects the villous enterocytes of the small intestine and causes them to slough. This in turn leads to malabsorption and dehydration, signs characteristic of transmissible gastroenteritis (TGE) (Saif and Bohl, 1986).

Porcine respiratory coronavirus (PRCV) is thought to be a respiratory variant of TGEV (Pensaert, 1989). It replicates in the respiratory tract of swine with little clinical disease, but some isolates have been found to cause lesions which may predispose the pig to secondary infections (Halbur et al., 1993). PRCV was first discovered in Belgium in 1986, and has now become widespread in Europe (Pensaert et al., 1986). The prevalence of PRCV in the United States is not known, but several researchers have isolated it from U.S. herds (Hill et al., 1989; Paul et al., 1992; Vaughn et al., 1995).

PRCV and TGEV have been shown to be antigenically and genetically related (Laude et al., 1993). TGEV and PRCV are genetically similar, and yet show differences in tissue tropism and pathogenicity which make them good candidates for the study of coronavirus genes which are involved in tissue tropism and virulence (Vaughn et al., 1995).

Coronavirus replication is unique among RNA viruses. They form a

3'-terminal nested set of subgenomic mRNAs which are polycistronic in nature, but monocistronic in function because only the unique 5' end of each mRNA is translated (Spaan 1988). TGEV and PRCV form seven or eight mRNAs during replication (Wesley et al., 1990). Their genome is similar to other coronaviruses. Prior to each ORF on the genome there is a conserved leader binding sequence CTAAAC which is believed to be used during the transcription of the mRNAs as a consensus leader binding site (Lai, 1990; Spaan, 1988).

The mRNA 1 encodes for the RNA-dependent RNA polymerase which is necessary for viral replication (Lai, 1990). The mRNA 2 encodes for the spike (S) protein which is responsible for viral attachment and tissue tropism (Lai, 1990). The mRNAs 3 and 3-1 encode for two nonstructural proteins, the fuction of which is unknown at this time. They have been hypothesized to be involved in pathogenicity (Wesley et al., 1990). A small integral membrane protein (sM) is encoded for by mRNA 4 (Godet et al., 1992). The membrane protein (M) and the nucleocapsid protein (N) are encoded by mRNAs 5 and 6 (Spaan, 1988). The last mRNA, mRNA 7, has been shown to encode for a small protein which has been implicated in binding DNA in infected cells (Garwes et al.,1989).

Studies which have compared the genomic sequences of PRCV to TGEV have shown two major differences between the two viruses. The S gene of PRCV, when compared to TGEV, has a large deletion which ranges from 621- to 681- nucleotides in length (Britton et al., 1991; Laude et al., 1993; Rasschaert et al., 1990; Vaughn et al., 1994). This is thought to play a role in determining tissue specificity (Sanchez et al., 1992). In all PRCV isolates there were alterations in the 3 and 3-1 genes either in the leader binding sequence or in the gene sequence itself. This leads to the mRNAs from these ORFs either not being produced, or producing mRNAs which code for truncated proteins (Laude et al., 1993; Page et al.,1991; Rasschaert et al., 1990; Vaughn et al., 1994). The 3 and 3-1 genes in TGEV have been hypothesized to be important in the

pathogenesis of TGEV (Wesley et al., 1989). The U.S. PRCV isolates AR310 and LEPP are the only ones which have been shown to have an intact 3 gene, all other isolates which have been characterized show an alteration or deletion in the 3 or 3-1 region (Vaughn et al. 1994).

In this study, the molecular characteristics of four newly cultured PRCV isolates were analyzed. The size and location of the 5' S gene deletion was examined. Genes 3 and 3-1 were also sequenced. The sequence of these isolates were then compared to four isolates which were studied previously by Vaughn et al. (1994), as well as the virulent Miller strain of TGEV.

The biological characteristics of all eight PRCV isolates used in this study were also examined. Variations in plaque sizes, growth patterns in two cell lines, and CPE differences were all utilized to compare these isolates. Some of the strains that have been previously characterized vary in pathogenicity (Halbur et al., 1994). By utilizing these isolates in our study we determined if there was a link between the differences seen *in vitro* and their relevance *in vivo*.

Materials and Methods

Cell Culture

Two cell lines were used in this study. The swine testicular (ST) cell line was used to propagate and isolate PRCV (McClurkin and Norman, 1966; Zhu et al., 1990). The ST cells were maintained in Eagle's minimum essential medium (MEM) (Gibco, Grand Island, NY) with 10 percent fetal bovine serum (FBS) (Gibco, Grand Island, NY), sodium bicarbonate (2.9 g/L) (Fisher Scientific, Fair Lawn, NJ) and lactalbumin enzymatic hydrolysate (5.0 g/L) (Sigma, St. Louis, MO). The ST cells were maintained in 75 cm^2 tissue culture flasks (Costar, Cambridge, MA) with 3-4 days between subculturing. They were grown in a 37°C incubator with 5% CO_2. The cells were subcultured by decanting off old

medium and washing three times with 10 ml of rinse saline. After decanting off the final wash, 3 ml of 0.2% trypsin was added and the flask was incubated at 37°C for 5 minutes. The flask was then tapped gently to remove the cells from the flask surface. The cells were resuspended in 100 ml of medium and subcultured at a 1:4 ratio.

The second cell line that was used in the study was a porcine intestinal epithelial cell line, IPEC-1. This cell line is a continuous line of epithelial cells derived from the intestine of a 12 hour old piglet. The cell line was kindly provided by Dr. H.M. Berschneider of North Carolina State University. The IPEC-1 cells were grown in DMEM/F12 medium (Gibco, Grand Island, NY) which was supplemented with epidermal growth factor (EGF) (1μg/L)(Collaborative Biomedical Products, Bedford, MA), ITS premix (Collaborative Biomedical Products, Bedford, MA) which contained insulin (5mg/L), transferrin (5mg/L) and selenium (5μg/L) and 5% fetal bovine serum (FBS) (Gibco, Grand Island, NY). The IPEC-1 cells were propagated in 75cm^2 flasks with 7-10 days between subcultures. They were maintained in a 37°C incubator with 5% CO_2. These cells were subcultured in much the same way as the ST cells. The old medium was poured off when the cells began to detach slightly from the flask (7-10 days). The monolayer was then rinsed three times with 10 ml of calcium-magnesium free phosphate buffered saline (CMF-PBS). The last rinse was decanted off and 3 ml of 0.1% trypsin was added to the flask. The cells were placed back in the incubator for 10-20 minutes to allow the cells to detach from the monolayer. The cells were resuspended in 75 ml of medium and subcultured at a 1:3 ratio.

Viruses

Eight PRCV isolates were used in this study. Four of the isolates; AR310, LEPP, IA1894 and ISU-1, have been described in a in previous study in which

they looked at the S gene sequences and pathogenicity (Vaughn et al., 1994). AR310 was isolated from intestinal homogenates of a piglet with TGEV in a swine herd in Arkansas in 1989 (Halbur et al., 1993). LEPP and IA1894 were isolated from nasal swabs which were collected from herds in Iowa with pneumonia in 1991 and 1992 (Vaughn et al . 1995). The isolate ISU-1 was obtained from Dr. Howard Hill of the Iowa State Veterinary Diagnostic Lab (Hill et al., 1989). The four new isolates were isolated from various sources. NVSL 5170 was isolated from a fecal sample received by the National Veterinary Services Laboratory in Ames, IA. The herd had a history of ongoing diarrhea, and 10% mortality among neonates. Low to moderate numbers of coronavirus like particles were visible in the sample when viewed under electron microscopy. Samples were found to be negative for TGEV by two diagnostic labs. The owner also had purchased pigs from the United Kingdom. Due to the difficulty in the detection of TGEV and history of the addition of pigs from Europe porcine epidemic diarrhea was suspected. However virus isolation and sequence analysis showed that it was not true. When tissues were examined using in situ hybridization, a weak positive signal for TGEV was seen and both TGEV and PRCV were isolated from a fecal sample (Halbur et al., 1995). Isolate 725 was from a herd in Iowa in 1992 where the nursery pigs had a mortality rate of approximately 10%. The nursery pigs had some sneezing and some respiratory illness prior to necropsy. Upon necropsy pneumonia-type lesions were seen. The pigs had high titers to TGEV, but did not show any enteric disease signs. The isolate 306 was also from a herd in Iowa which had lung lesions and signs of pneumonia upon necropsy. This herd was also positive for porcine respiratory and reproductive syndrome (PRRSV). Isolate PON was obtained from a herd in Nebraska which had clinical respiratory signs. Upon necropsy evidence of bronchopneumonia was present. They had high titers to TGEV with

no enteric disease symptoms which is indicative of PRCV infection. PRRSV was also isolated from these pigs.

All of the viruses were isolated and propagated in ST cells. They were plaque purified three times before being used for further characterization. Stock virus was prepared from these isolates and titrated in ST cells then stored at -70°C. The titer was expressed as tissue culture infectious dose ($TCID_{50}$).

Growth Curve

A two step growth curve was done to determine the growth rate and titer of each isolate in the ST cell line and compare it to an intestinal epithelial cell line, IPEC-1. Twelve well plates (Costar, Cambridge, MA) containing monolayers of ST cells or IPEC-1 cells were inoculated with 0.1 MOI of each isolate in MEM or DMEM/F12 respectively. After one hour, the virus inoculum was removed and 2 ml of MEM with 2% FBS or DMEM/F12 with 2% FBS was added to each well. At various times post-infection the wells were scraped with a cell scraper, the fluid was removed and placed in a 1.5 ml microcentrifuge tube, then stored at -70°C. When samples at all of the time points had been collected they were subjected to three cycles of freezing and thawing. Serial dilutions of each isolate was made and 96 well plates of three-day-old ST cells were inoculated to determine the virus titer. All time points of 8 isolates were done in duplicate.

Plaque Assay

Six-well plates of confluent ST cells were inoculated with 0.5 ml of each virus isolate. After a one hour incubation period at 37°C, the virus inoculum was removed. A 2ml agarose overlay was then added to each well. The agarose was composed of 2x Basal Minimal Essential Medium (BME) and 1.8% SeaPlaque agarose (FMC Bioproducts) in equal amounts with 1.5ml/100ml of

7.5% sodium bicarbonate and 1.6ml/100ml of 0.1% neutral red. The plates were allowed to sit for 15 minutes under the hood to allow the agarose to harden and they were then placed upside down in the incubator at 37°C and allowed to incubate for four days. The plaques were then observed on a light box and the diameter was measured. One hundred plaques were measured for each isolate.

CPE Differences

Monolayers of ST and IPEC-1 cells were grown in 25 cm^2 tissue culture flasks. They were infected with 0.1 MOI of virus and observed every 24 hours for CPE. Differences were recorded and observations were carried out for 96 hours post infection. This was done for each virus isolate.

RNA Isolation

ST cells were infected with 0.1 MOI with AR310, ISU-1, IA1894, LEPP, PON, 5170 NVSL, 306, or 725. At 48 hours post infection, the medium was removed and total RNA was isolated from the infected ST cells by using a micro RNA isolation kit (Stratagene, La Jolla, CA). The RNA was washed with 75% ethanol, dissolved in DEPC-treated water, and stored at -70°C. The concentration and quality of RNA was determined by OD at 260 and 280nm.

cDNA Synthesis and PCR Amplification

First strand cDNA synthesis was done on total RNA from infected ST cells, using the cDNA Cycle Kit (Invitrogen, San Diego, CA) with random primers. We have found that for the 3 and 3-1 genes we get better results with specific primers 538 and 622 (Table 2) for cDNA synthesis, so we have used those for that portion of the experiment.

The cDNA was amplified by PCR using TaqDNA polymerase (Boehringer

Table 2. Specific primers used for S gene and 3, 3-1 gene PCR

Primer Description	Sequence	Orientation	Gene
184	GTAAAAACATTAGCCACATA	Reverse	S
185	AGGGTAAGTTGCTCATTAG	Forward	S
538[a]	CTATTGAAAAAGTGCACGTC	Forward	3, 3-1
622[a]	AATGATGCTAATGACCATTC	Reverse	3, 3-1

[a] from Vaughn et al. (1995).

Mannheim, Indianapolis, IN) and by using specific primers (Table 2). The 5'-half of the S gene of the PRCV isolates was amplified with the primers 184 and 185 using the following PCR conditions: 35 cycles of 30 seconds at 92°C, 30 seconds at 49°C, and 45 seconds at 72°C in a DNA thermal cycler.

The 3 and 3-1 region was amplified with the primers 538 and 622 using the following conditions: 1 cycle of 1 minute at 94°C, 1 minute at 49°C, and 5 minutes at 72°C; 30 cycles of 1 minute at 94°C, 1 minute at 49°C, and 1.5 minutes at 72°C; followed by 1 cycle of 1 minute at 94°, 1 minute at 49°C and 5 minutes at 72°C in a DNA thermal cycler.

PCR Product Purification

One-tenth of the PCR product was run on a 1% agarose gel to evaluate the success of the PCR. The rest of the product was purified directly from the PCR reaction by using the Wizard PCR Preps DNA Purification System for rapid purification of DNA fragments (Promega, Madison, WI). Two microliters of the purified product was run on a 1% agarose gel to determine the product concentration.

Sequence Analysis

The RT-PCR product for the S gene and the 3, 3-1 genes of each isolate was purified and sequenced directly with an automated DNA sequencer (Applied Biosystem) using virus-specific primers. The sequence of the S gene was determined with primers 184 (5' GTAAAAACATTAGCCACATA 3') and 185 (5' AGGGTAAGTTGCTCATTAG 3'). The sequence of the 3, and 3-1 genes were determined with primers 538 (5' CTATTGAAAAAGTGCACGTC 3') and 622 (5' AATGATGCTAATGACCATTC 3') and also internal primers 048 (5' GCATAGGTCCTAAAAGTGTCATTG 3') and 118 (5' TTTGTGTGTTTACTTCTTCA 3'). Sequence alignment and analysis was done using the Gene Works and Mac Vector programs for Macintosh.

Results

Growth Curve

The growth curves for all virus isolates were similar in the IPEC-1 cells. The peak tiers were in the range of 10^9 $TCID_{50}$/ml (Table 3). Isolates AR310, LEPP, PON, 5170 NVSL, and 306 peaked at 24 hours post infection. IA1894, ISU-1, and 725 did not peak until 48 hours post infection. All isolates had an initial peak between 24 and 30 hours post infection, followed by a slight drop in titer, then a second peak between 42 and 48 hours post infection (Figure 1). The titers were lower in the ST cells (Figure 2). They all reached a titer between 10^7 to 10^8 $TCID_{50}$/ml. The titers were at their peak when the CPE was approximately 30-50%, which was around 24-30 hours post infection. All isolates reached 100% CPE by 84 hours post infection.

The titers reached their peak before CPE became extensive. The CPE took longer to develop in the IPEC-1 cells, not becoming extensive until 42 hours post infection. The isolate IA1894 never caused any CPE in the IPEC-1 cells,

Table 3. Maximum Titer ($TCID_{50}$) for each PRCV isolate

PRCV Isolate	ST cells	Hours Post Infection	IPEC-1 Cells	Hours Post Infection
AR310	$4.2x10^8$	30	$3.1x10^9$	24
LEPP	$1.0x10^8$	30	$3.1x10^9$	24
IA1894	$7.4x10^7$	42	$2.3x10^9$	48
ISU-1	$5.6x10^8$	24	$3.1x10^9$	48
PON	$7.4x10^7$	30	$3.1x10^9$	24
725	$3.1x10^8$	48	$3.1x10^9$	48
5170 NVSL	$3.1x10^7$	48	$3.1x10^9$	24
306	$1.0x10^8$	60	$3.1x10^9$	30

but replicated to high titers, indicating that it may be capable of developing a persistent infection of the cell line.

Plaque assay

There was variation in the plaque sizes among all of the isolates. The isolate 725 developed the largest plaques with an average size of 5.255 mm (Table 4). IA1894 had the smallest plaque size of all the isolates with the average size being 1.04mm. There doesn't seem to be a correlation between virulence and plaque size, as both LEPP and AR310 are more virulent in swine than IA1894, but LEPP has the second smallest plaque size of all the isolates studied. Further studies could be done in this area to explore the possibility of a correlation between the plaque size and virulence.

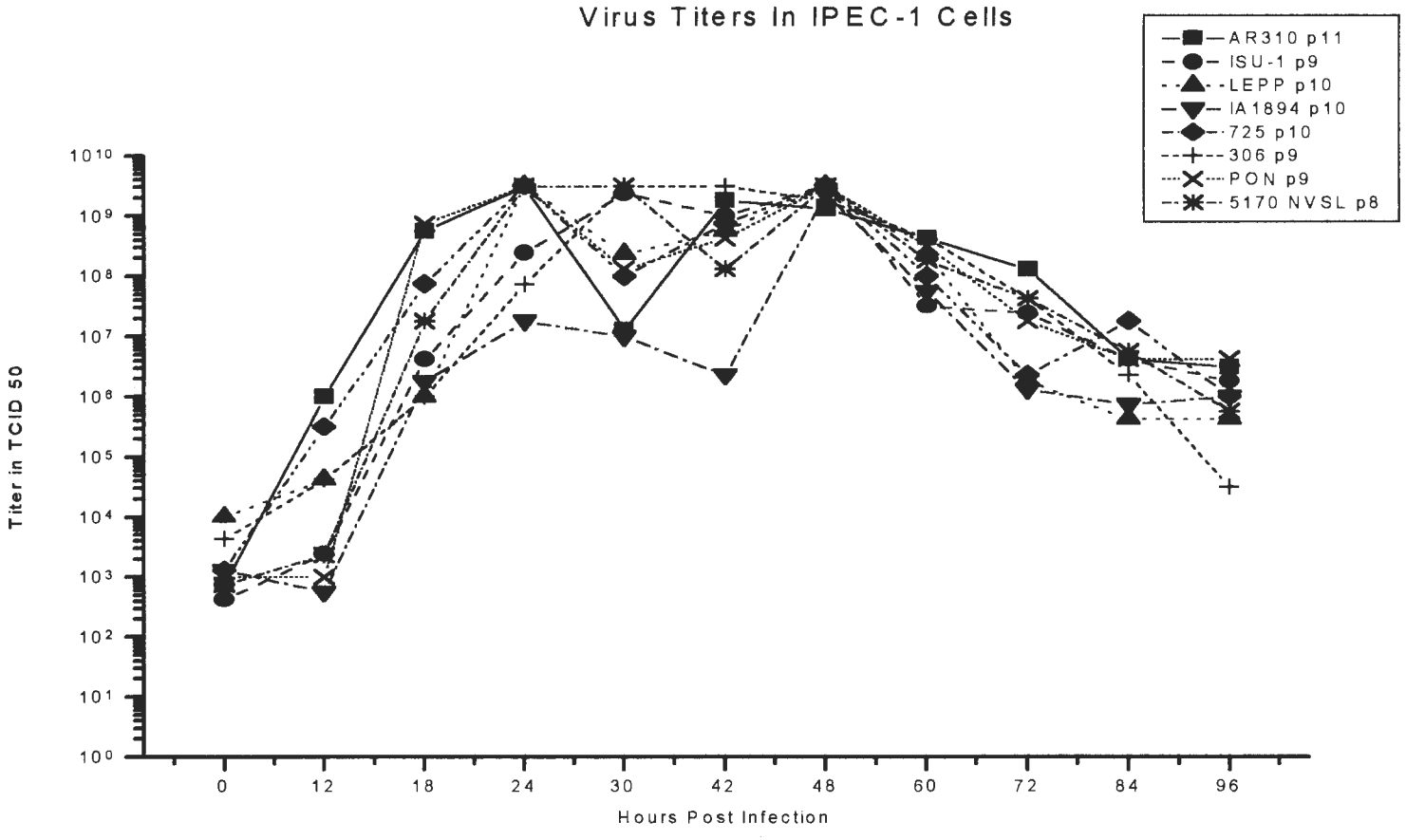

Figure 1. This figure shows the titer of PRCV isolates in IPEC-1 cells at various timepoints. Titers were measured in $TCID_{50}$. All timepoints were done in duplicate.

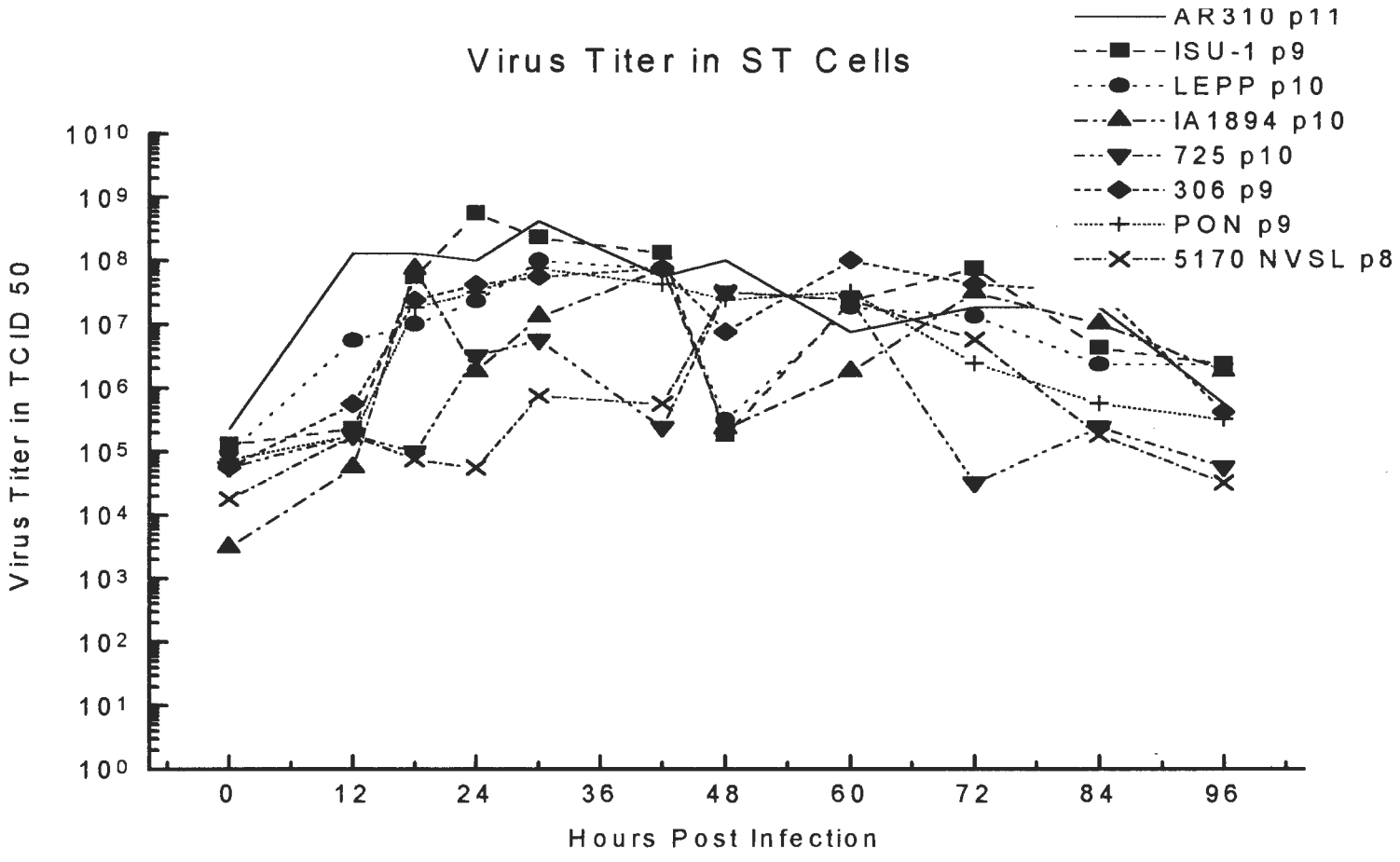

Figure 2. This figure shows the titer of PRCV isolates in ST cells at various timepoints. Titers were measured in $TCID_{50}$. All timepoints were done in duplicate.

Sequence analysis of the S, 3, and 3-1 genes

Analysis of the sequences shows that there are differences in deletions observed amongst various isolates. The isolates IA1894, ISU-1, AR310, and LEPP were sequenced previously (Vaughn et al., 1995). Isolate 725 had the same deletion as AR310 and LEPP of 621 base pairs located 44 base pairs from the S gene start site (see Figure 3). It also has an intact gene 3 and a 1 nucleotide deletion in the gene 3-1 at the same place as AR310 and LEPP. Isolate 306 has a 675 base pair deletion in the S gene which starts 54 base pairs from the start codon. This deletion is unique among the isolates sequenced so far. It has an intact leader binding region in the 3 gene, but an insertion of 29 base pairs codes for a stop site and causes the protein to be truncated. There is no leader binding sequence in the gene 3-1 and no start site (see figure 4). Instead there is a 52 bp deletion in this area, so the gene 3-1 is not transcribed. The isolate NVSL5170 has the longest deletion in the S gene sequenced so far. The deletion is 711 base pairs and begins only 27 base pairs from the start site. It has a normal gene 3, the same as AR310, LEPP and 725, but the gene 3-1 is truncated. The isolate PON has a 681 base pair deletion, 62 base pairs from the S gene start site. Although the size is the same, it begins 3 base pairs further downstream than the ISU-1 deletion. Both the genes 3 and 3-1 are truncated in this isolate.

Discussion

The results of our experiments show that there is a wide diversity among isolates of PRCV in the United States. By characterizing a few isolates from the Midwestern part of the US we have seen that there are large variations among the S, 3 and 3-1 genes of these isolates. All isolates characterized in Europe had similar deletions in both the S, 3, and 3-1 genes, while many of our isolates

had deletions specific for them. It has been speculated that the 3 and 3-1 genes are involved in the virulence of TGEV and PRCV. Further studies need to be done with these isolates to determine if this is a valid conclusion. We have found two new isolates with intact 3 genes, and one isolate with no leader binding sequence or start site in the 3-1 gene, so therefore it may make no product for that region.

The fact that the isolates replicate to a higher titer in the IPEC-1 cells is interesting, because they are derived from intestinal epithelial cells. It has been documented previously that there is little replication of PRCV in the gut (Cox et al., 1990), yet they grow to a higher titer in this intestinal cell line than in the ST cells which are routinely used to propagate PRCV. This leads to the conclusion that some replication of PRCV can occur in the intestine, but it must be remembered that this is a continuous cell line, and may no longer have all the same qualities of the cells *in vivo*. It was interesting that the isolate IA1894 caused no CPE in the IPEC-1 cells. This aspect should be studied more closely to determine if there is some potential for vaccine development with this isolate.

Our studies have confirmed the results of other researchers that PRCV isolates have a deletion in the S gene, and are highly variable in the 3 and 3-1 genes. Further studies need to be done to determine the function of genes 3 and 3-1 and to define the roles that they play in pathogenicity.

Table 4. Measurement of plaque size diameter for each isolate[a]

Plaque #	IA1894	306	LEPP	AR310	ISU-1	5170NVSL	PON	725
1	1.5	3.5	2.5	3.5	5.0	4.0	5.0	5.5
2	1.2	4.0	2.0	3.0	4.5	3.5	4.0	5.5
3	1.2	3.0	2.0	2.5	4.0	3.5	4.0	4.5
4	1.0	3.5	2.0	3.0	4.0	4.0	3.5	4.5
5	1.2	4.0	1.0	3.5	4.5	3.5	4.0	4.5
6	1.0	3.5	1.5	2.5	3.5	3.5	4.0	6.0
7	1.5	3.0	2.5	4.0	5.0	4.0	4.0	5.5
8	0.8	4.0	2.0	3.0	5.0	4.0	4.0	6.0
9	1.5	3.0	1.5	3.0	4.5	4.5	3.5	6.0
10	1.0	2.5	1.5	3.5	4.0	5.0	5.0	4.0
11	1.5	4.0	2.0	4.5	5.0	4.0	4.5	5.0
12	1.0	4.5	2.0	4.0	4.5	4.0	4.5	5.5
13	0.8	3.0	1.5	4.0	4.0	4.0	3.0	4.0
14	1.0	3.5	2.5	2.5	3.5	3.0	3.5	4.5
15	0.8	2.5	1.0	3.0	4.0	4.5	4.0	6.0
16	1.5	3.0	2.0	3.5	5.0	5.0	4.0	5.0
17	2.0	2.5	2.0	3.0	5.0	4.0	4.5	5.5
18	1.0	3.5	1.0	4.0	4.0	4.0	5.0	4.5
19	0.8	4.0	1.5	3.0	4.5	5.0	5.0	4.5
20	1.0	4.0	2.0	4.0	4.0	4.0	4.0	4.5
21	0.5	3.0	2.5	3.0	5.5	5.0	4.0	6.0
22	1.0	3.0	2.0	3.5	3.5	3.5	3.5	6.0
23	1.0	3.5	1.0	2.5	4.0	4.5	3.0	5.0
24	1.0	3.0	1.5	2.0	5.0	4.0	3.5	6.5
25	1.0	2.5	1.5	3.0	5.0	4.0	4.0	5.5
26	1.0	3.0	1.5	4.0	4.5	5.0	4.0	5.5
27	1.0	3.5	2.0	3.0	5.5	4.5	4.5	4.0
28	1.5	4.0	2.0	2.5	4.0	3.5	5.0	4.0
29	1.0	4.5	2.5	4.0	4.5	4.0	4.5	4.5
30	1.0	3.5	2.5	4.0	4.0	4.0	4.0	5.5
31	1.0	3.0	1.0	3.0	4.0	4.0	3.5	6.0
32	1.2	4.0	2.0	3.5	4.5	3.0	4.0	6.0
33	1.5	4.0	2.0	2.5	3.5	4.5	4.0	5.0
34	1.0	2.5	2.5	2.5	3.0	4.0	3.5	5.0
35	0.5	4.0	1.0	3.0	5.0	5.0	5.0	4.5
36	1.0	4.5	1.5	3.0	3.5	5.0	5.0	5.0
37	1.0	3.0	1.5	3.0	4.0	4.5	4.0	5.5
38	0.8	3.5	1.0	4.0	5.0	3.5	4.5	4.0
39	0.8	3.0	2.5	4.5	5.0	3.5	4.0	6.0
40	0.5	4.0	2.0	3.0	4.0	4.5	4.5	6.0
41	1.5	3.5	2.0	3.5	3.0	4.0	4.5	4.5
42	0.5	4.0	1.5	3.5	3.5	3.5	4.0	4.5
43	0.8	3.0	2.0	2.5	4.0	5.0	3.5	6.0
44	1.5	3.0	2.5	2.5	4.5	4.0	5.0	5.5

[a] Measurements were taken in mm

Table 4.(continued)

Plaque #	IA1894	306	LEPP	AR310	ISU-1	5170NVSL	PON	725
45	1.0	2.5	1.0	4.0	4.0	4.0	3.5	5.0
46	1.0	2.5	1.5	4.0	5.5	4.5	4.0	5.0
47	0.5	3.0	2.0	3.0	4.0	3.5	4.5	4.5
48	0.8	3.5	2.0	3.5	4.5	3.0	4.0	6.0
49	1.0	3.0	1.5	4.0	5.0	4.0	4.0	5.5
50	1.5	2.5	1.5	4.5	5.0	3.5	5.0	6.0
51	1.5	4.0	1.0	4.0	4.5	4.5	4.0	4.5
52	1.0	3.5	2.5	3.0	4.0	4.0	4.5	4.0
53	1.0	3.0	1.0	4.0	4.5	4.0	3.5	4.5
54	1.0	3.0	1.5	2.5	3.5	5.0	4.0	6.0
55	0.8	2.5	2.0	3.0	3.5	4.5	4.0	5.0
56	1.0	2.5	2.0	3.0	5.0	4.0	5.0	6.0
57	0.8	2.5	2.0	3.5	4.5	3.0	5.0	5.5
58	1.2	3.0	1.8	2.5	4.0	4.5	4.0	5.0
59	0.5	3.0	2.0	2.0	4.5	5.0	4.0	5.5
60	0.8	3.5	2.0	3.0	5.0	4.0	5.0	5.0
61	1.0	4.0	2.0	3.0	5.0	4.0	5.0	6.0
62	1.2	4.5	2.0	3.5	4.5	5.0	3.5	5.5
63	1.2	3.5	1.5	3.5	5.5	4.5	5.0	6.0
64	1.5	3.0	2.5	3.0	4.0	3.5	5.0	5.0
65	1.2	4.0	2.5	4.0	4.5	3.5	3.5	6.0
66	1.0	3.0	1.5	4.0	4.0	4.0	3.5	5.5
67	0.8	3.5	1.5	3.5	4.5	3.5	4.0	6.0
68	1.0	3.0	1.5	2.5	5.0	4.5	4.0	5.0
69	1.0	2.5	1.0	2.5	3.5	4.0	4.0	5.5
70	1.2	3.0	2.0	3.0	3.5	5.0	4.0	6.0
71	1.2	3.5	1.0	3.0	4.0	5.0	2.5	6.0
72	1.0	4.0	2.5	3.5	4.5	4.0	4.0	5.0
73	1.0	3.5	2.0	4.0	5.0	4.0	4.0	5.5
74	1.0	3.0	2.0	3.5	3.0	4.5	4.0	5.0
75	1.2	4.0	1.0	3.5	5.0	3.5	4.0	6.0
76	1.0	4.0	1.5	3.0	4.0	3.0	5.0	5.5
77	1.0	3.0	2.0	3.0	4.0	4.0	4.0	4.5
78	1.0	2.5	2.0	2.5	4.5	3.5	4.0	5.0
79	1.0	3.0	2.0	4.0	5.0	4.5	3.5	6.0
80	1.0	3.5	2.0	3.5	4.0	5.0	3.5	6.0
81	1.0	3.0	1.5	4.0	4.0	4.0	4.0	4.5
82	1.0	2.5	1.5	4.5	4.5	4.0	4.0	5.5
83	1.0	3.5	2.0	4.0	4.0	4.0	3.5	5.0
84	1.2	4.0	2.0	3.0	5.0	4.5	5.0	5.0
85	1.2	4.0	2.0	3.0	5.0	3.5	3.0	6.0
86	1.2	4.5	2.0	3.0	4.0	5.0	3.5	6.0
87	1.0	2.5	2.5	2.5	3.5	4.0	4.0	6.0
88	1.0	4.0	2.5	2.5	4.0	4.0	4.0	5.5
89	0.8	3.0	2.5	3.0	4.5	4.0	5.0	4.0
90	1.0	2.5	1.5	3.5	5.0	3.5	5.0	4.5
91	1.0	2.5	1.0	3.0	3.0	4.5	3.0	5.0

Table 4.(continued)

Plaque #	IA1894	306	LEPP	AR310	ISU-1	5170NVSL	PON	725
92	1.0	3.5	0.8	3.0	3.5	5.0	4.0	5.0
93	0.8	4.0	1.0	4.0	4.0	4.0	4.0	5.5
94	1.0	4.0	1.5	4.0	4.5	3.0	4.0	6.0
95	0.8	4.5	1.5	3.0	4.0	4.5	4.0	5.0
96	1.2	2.5	1.5	2.5	5.5	4.0	3.5	4.5
97	1.5	4.0	1.5	4.0	3.5	4.0	4.0	4.0
98	1.2	3.0	1.5	4.0	4.0	5.0	3.5	5.5
99	0.8	3.0	2.0	3.0	4.5	4.5	3.0	6.0
100	1.0	4.0	1.5	3.0	5.0	4.0	5.0	6.0
Average	1.0	3.3	1.8	3.3	4.3	4.1	4.1	5.2
Standard Deviation	0.3	0.6	0.5	0.6	0.6	0.6	0.6	0.7

```
          S Start
          +1→
TGEV      ATGAAAAAAT TATTTGTGGT TTTGGTTGTA ATGCCATTGA TTTATGGAGA     50
ISU-1     ATGAAAACAT TATTTGTGGT TTTGGTTATA ATGCCATTGA TTTATGGAGA     47
AR310     ATGAAAACAT TATTTGTGGT TTTGGTTATA ATGCCATTGA TTTATGG---     47
LEPP      ATGAAAACAT TATTTGTGGT TTTGGTTATA ATGCCATTGA TTTATGG---     47
1894      ATGAAAACAT TATTTGTGGT TTTGGTTATA ATGCCATTGA TTTA------     44
725       ATGAAAACAT TATTTGTGGT TTTGGTTATA ATGCCATTGA TTTATGG---     47
306       ATGAAAAAAT TATTTGTGGT TTTGGTCTTA ATGCCATTGA TTTATGGAGA     50
5170      ATGAAAACAC TATTTGTGGT TTTGGTT--- ---------- ----------     27
PON       ATGAAACCAT TATTTGTGGT TTTGTTTATA ATGCCATTGA TTTATGGAGA     50
```


Figure 3. This is the alignment of the nucleotide sequences from the 5'end of the S gene of PRCV isolates 725, 306, 5170, and PON as compared to TGEV isolate CHV and PRCV isolates ISU-1, AR310, LEPP, and 1894. The PRCV isolate 725 has an identical deletion to isolates AR310 and LEPP of 621 nucleotides. The isolate 306 has a 675 nucleotide deletion which is unique to this isolate. Isolate 5170 has the largest deletion described so far with a size of 711 nucleotides. PON has a deletion of 681 nucleotides which is identical to isolate ISU-1. The start codon of the S gene is marked with +1→. The nucleotides which make up the predicted signal peptide region of the TGEV S gene as noted by Rasschaert and Laude (1987) are underlined. The nucleotide sequences of PRCV isolates 725, 306, 5170, and PON were determined in this study. The other nucleotide sequences were determined by Vaughn et al., 1995.

```
TGEV      CAATTTTCCT TGTTCTAAAT TGACTAATAG AACTATAGGT AACCATTGGA     100
ISU-1     TAATTTTCCT TG-------- ---------- ---------- ----------     62
AR310     ---------- ---------- ---------- ---------- ----------     47
LEPP      ---------- ---------- ---------- ---------- ----------     47
1894      ---------- ---------- ---------- ---------- ----------     44
725       ---------- ---------- ---------- ---------- ----------     47
306       CAATTTT--- ---------- ---------- ---------- ----------     57
5170      ---------- ---------- ---------- ---------- ----------     27
PON       TAATTTTCCT TG-------- ---------- ---------- ----------     62

TGEV      AGCCGCTGGC ACGCTTGTAG ACCTTTGGTG GTTTAATCCT GTTTATGATG     700
ISU-1     ---------- ---------- ---------- ---------- ----------     62
AR310     ---------- --------AG ACCTTTGGTG GTTTAATCCT GTTTATGATG     79
LEPP      ---------- --------AG ACCTTTGGNG GTTTAATCCT GTTTATGATG     79
1894      ---------- ---------- ---------- ---------- ----------     44
725       ---------- --------AG ACCTTTGGTG GTTTAATCCT GTTTATGATG     79
306       ---------- ---------- ---------- ---------- ----------     57
5170      ---------- ---------- ---------- ---------- ----------     27
PON       ---------- ---------- ---------- ---------- ----------     62

TGEV      TCAGTTATTA TAGAGTTAAT AATAAAAATG GTACTACCGT AGTTTCCAAT     750
ISU-1     ---------- ---------- ---------- ---------- ---TTCCAAT     69
AR310     TCAGTTATTA TAGAGTTAAT AATAAAAATG GTACTACCGT AGTTTCCAAT     129
LEPP      TCAGTTATTA TAGAGTTAAT AGTAAAAATG GTACTACCGG AGTTTCCAAT     129
1894      ---------- ---------- ----AAAATG GTACTACCGT AGTTTCCAAT     70
725       TCAGTTATTA TAGAGTTAAT AATAAAAATG GTACTACCGT AGTTTCCAAT     129
306       ---------- ---------- ----ACAATT TTACTACCGT AGTTTCCAAT     83
5170      ---------- ---------- ---------- --------GT AGTTTCCAAT     39
PON       ---------- ---------- ---------- ---------- ---TTCCAAT     69
```

Figure 3.(continued)

```
           Primer  538                S
           ========== ==========     Stop
LEPP       CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTTTATC---   47
AR310      CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTTTATC---   47
ISU-1      CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAG TTTTATC---   47
1894       CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTTTATC---.  47
725        CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTTTATC---   47
306        CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTCTATCATC   50
5170       CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTTTATC---   47
PON        CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTTTATC---   47
CHV        CTATTGAAAA AGTGCACGTC CATTAAATTT AAAATGTTAA TTTTATTATC   50
           ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^  ^^ ^^^
```

Figure 4. This figure shows the comparison of the nucleotide sequence of gene 3 and 3-1 regions of the PRCV isolates with each other and the TGEV isolate CHV. Positions of the leader binding regions are marked with (*****). The positions of the primers 538, 118, 048, and 622 are marked with (=====). The start codons of the genes are marked with +1→, and the stop codons are underlined. Positions having identical nucleotides are marked with (^), and positions of deleted nucleotides are marked with (-----). The nucleotide sequences of the PRCV isolates 725, 306, 5170, and PON were determined in this study. The nucleotide sequence of other PRCV and the TGEV isolate CHV was determined previously by Vaughn et al., 1995.

```
LEPP      TGCTATAATA GCATTTGTT- ---------- -A-----TTA AGGATGATGA   80
AR310     TGCTATAATA GCATTTGTT- ---------- -A-----TTA AGGATGATGA   80
ISU-1     CGCTAT---- ---------- ---------- ---------- ----------   53
1894      TGCTATAATA GCATTTGTT- ---------- -A-----TTA AGGATGATGA   80
725       TGCTATAATA GCATTTGTT- ---------- -A-----TTA AGGATGATGA   80
306       TGCTATAATA GCAGTTGTTT CTGCTAGAAG AATTTTGTTA AGGATGATGA   100
5170      TGCTATAATA GCATTTGTT- ---------- -A-----TTA AGGATGATGA   80.
PON       TGCTATAATA GCATTTGTT- ---------- -A-----TTA AGGATGATGA   80
CHV       TGCTATAATA GCATTTGTT- ---------- ------GTTA AGGATGATGA   83
           ^^^^^
```

```
                              *** ****
LEPP      ATAAAGTCC- TTAAGAACTA AACTTTCTGG TCATTACAG- ----------   118
AR310     ATAAAGTCC- TTAAGAACTA AACTTTCAGG TCATTACAG- ----------   118
ISU-1     ---------- ---------- ---------- ---------- ----------   53
1894      ATAAAGTCC- TTAAGAACTA AACTTTCAGG TCATTACAG- ----------   118
725       ATAAAGTCC- TTAAGAACTA AACTTTCAGG TCATTACAG- ----------   118
306       ATAAAGTCT- TTAAGAACTA AACTTACGAG TCATTACAGC AAAGCAAGGT   149
5170      ATAAAGTCC- TTAAGAACTA AACTTTCAGG TCATTACAG- ----------   118
PON       ATAAAGTCCA ACTCGAACTA AACTTTCAGG TCATTACAG- ----------   119
CHV       ATAAAGTCC- TTAAGAACTA AACTTTCGAG TCATTACAG- ----------   121
```

Figure 4. (continued)

```
                ORF 3
                +1→
LEPP      -GTCCTGTAT GGACATTGTC AAATCTATTA ATACATCCGT GGATGCTGTA   167
AR310     -GTCCTGTAT GGACATTGTC AAATCTATTA ATACATCCGT GGATGCTGTA   167
ISU-1     ---------- ---------- ---------- ---------- ----------    53
1894      -GTCCTGTAT GGACATTGGC AAATCCATTA TTACATCCGT GGATGCTGTA   167
725       -GTCCTGTAT GGACATTGTC AAATCTATTA ATACATCCGT GGATGCTGTA   167
306       TGTCCTGTAT GGACATTGTC AAGTCCATTT ACACATCCGT AGATGCTGTA   199
5170      -GTCCTGTAT GGACATTGTC AAATCCATTA ATACATCCGT GGATGCTGTA   167
PON       -GTCCTTTAT GGACATTGTC AAATCCATTA ATACATCCTT GGATGCTGTA   168
CHV       -GTCCTGTAT GGACATTGTC AAATCCATTA ATACATCCGT AGATGCTGTA   170


LEPP      CTTGACGAAC TTGATTGTGC ATACTTCGCT GTTACTCTTA AAGTAGAATT   217
AR310     CTTGACGAAC TTGATTGTGC ATACTTCGCT GTTACTCTTA AAGTAGAATT   217
ISU-1     ---------- ---------- ---------- ---------- ----------    53
1894      CTTGACGAAC TTGATTGTGC ATACTTCGCT GTAACTCTTA AAGTAGAATT   217
725       CTTGACGAAC TTGATTGTGC ATACTTCGCT GTTACTCTTA AAGTAGAATT   217
306       CTTGACGAAC TTGATTGTGC ATACTTTGCT GTAACTCTTA AAGTAGAATT   249
5170      CTTGACGAAC TTGATTGTGC ATACTTCGCT GTTACTCTTA AAGTAGAATT   217
PON       CTTGACTGAC TTGATTGTGC ATACTTCGCT GTAACTCTTA AGGTAGAATT   218
CHV       CTTGACGAAC TTGATTGTGC ATACTTTGCT GTAACTCTTA AAGTAGAATT   220
```

Figure 4.(continued)

```
LEPP       TAAGACTGGT AAATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   267
AR310      TAAGACTGGT AAATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   267
725        TAAGACTGGT AAATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   267
ISU-1      ----ACTGGT AAATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   99
1894       TAAGACTGGT AAATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   267
306        TAAGACTGGT AAATTACTTG TGTGTATAGG TTTTGGTTAC ACACTTCTTG   299
5170       TAAGACTGGT AGATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   267
PON        TAAGACTGAT AAATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   268
CHV        TAAGACTGGT AAATTACTTG TGTGTATAGG TTTTGGTGAC ACACTTCTTG   270
               ^^^^ ^ ^ ^^^^^^^^ ^^^^^^^^^^ ^^^^^^^ ^^ ^^^^^^^^^^
```

Primer 118

```
                                                        =======
LEPP       CGGCTAGGGA TAAAGCATAT GCTAAGCTTG GTCTCGCCAC TATTGAAGAA   317
AR310      CGGCTAGGGA TAAAGCATAT GCTAAGCTTG GTCTCGCCAC TATTGAAGAA   317
ISU-1      CGGCTAGGGA TAAAGCA--- --TAAGCTTG GTCTCGCCAC TATTGAAGAA   144
1894       CGGCTAGGGG TA-------- ---AAG---- -------CA- TATTGAAGAA   294
725        CGGCTAGGGA TAAAGCATAT GCTAAGCTTG GTCTCGCCAC TATTGAAGAA   317
306        CTGCTAAGGA TA-------T GCTAAGCTTG GTCTCTCCAT TATTGAAGAA   342
5170       CGGCTAGGGA TAAAGCATAT GCTAAGCTTG GTCTCGCCAC TATTGAAGAA   317
PON        ATGCTAGGGA TAAAGCATA- ----AGCTTG GTCTCGCCAC TATTGAAGAA   313
CHV        CGGCTAGGGA TAAAGCATAT GCTAAGCTTG GTCTCTCCAT TATTGAAGAA   320
             ^^^^ ^^  ^^            ^^            ^^  ^^^^^^^^^^
```

Figure 4. (continued)

```
                                                Primer  118
                                                =============
LEPP        G--------- ---------- ---------- TAAACACACA AAATCCAAAG   338
AR310       G--------- ---------- ---------- TAAACACACA AAATCCAAAG   338
ISU-1       G--------- ---------- ---------- TAAACACACA AAATCCAAAG   165
1894        G--------- ---------- ---------- TAAACACACA AAATCCAAAG   315
725         G--------- ---------- ---------- TAAACACACA AAATCCAAAG   338
306         GTCAATAGTC ATATAGTTGT TTAATATCAT TAAACACACA AAACCCAAAG   392
5170        G--------- ---------- ---------- TAAACACACA AAATCCAAAG   338
PON         G--------- ---------- ---------- TAAACACACA AAATCCAAAG   334
CHV         GT-------- ---------- ---------- -AAACACACA AAATCCAAAG   341
            ^                                 ^^^^^^^^^^ ^^^ ^^^^^^


LEPP        CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTATAGAAA AACTGTCATT   388
AR310       CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTATAGAAA AACTGTCATT   388
ISU-1       CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTATAGAAA AACTGTCATT   215
1894        CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTGTAGAAA AACTGTCATT   365
725         CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTATAGAAA AACTGTCGAA   388
306         TGTTAAGTGT TACAAAACAA TTAAA----- ---------- ----------   417
5170        CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTATAGAAA AACTGTCATT   388
PON         CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTATAGAAA AACTGTCATT   384
CHV         CATTAAGTGT TACAAAACAA TTAAAGAGAG ATTATAGAAA AACTGTCATT   391
              ^^^^^^^^ ^^^^^^^^^^ ^^^^^
```

Figure 4.(continued)

```
                          ORF  3-1
          ******           +1➔
LEPP      CTAAACTTTG TGTGAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   438
AR310     CTAAACTTTG TGTGAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   438
ISU-1     CTAAACTTTG TGTGAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   265
1894      CTAA-----G TGTTAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   410
725       CTAAACTTTG TGTGAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   438
306       ---------- ---------- -------GAC TTTTTCTTAG TACTCTGAGT   440
5170      CTAAACTTTG TGTGAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   438
PON       CTAAACTT-- TGTGAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   432
CHV       CTAAACTTCA TGCGAAAATG ATTGGTGGAC TTTTTCTTAA TACTCTGAGT   441
                                        ^^^ ^^^^^^^^^  ^^^^^^^^^^


LEPP      TTGGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   488
AR310     TTGGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   488
ISU-1     TTTGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   315
1894      TTTGTAATTG TTAGTAACCA TTCTATTGTT AATAATACAG CAAATGTGCA   460
725       TTTGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   488
306       TTTGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   490
5170      TTTGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   488
PON       TTTGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   482
CHV       TTTGTAATTG TTAGTAACCA TTCTATTGTT AATAACACAG CAAATGTGCA   491
          ^^ ^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^ ^^^^ ^^^^^^^^^^
```

Figure 4.(continued)

```
LEPP      TCATACACAA CAAGACCGTG TTATAGTACA ACAGCATCAG GTTGTTAGTG   538
AR310     TCATACACAA CAAGACCGTG TTATAGTACA ACAGCATCAG GTTGTTAGTG   538
ISU-1     TCATACACAA CAAG------ ---------- ---------- ----------   329
1894      TCATACACAA CAAGACCGTG TTATAGTACA ACATCATCAG GTTATTAGTG   510
725       TCATACACAA CAAGACCGTG TTATAGTACA ACAGCATCAG GTTGTTAGTG   538
306       TCATATACAA CAAAAACGTG TTATAGTACA ACAGCATCAG GTTGTTAGTG   540
5170      CCACACACAA CAAGACCGTG TTATAGTATA ACAGCATCAG GTTGTTAGTG   538
PON       TCATACACAA CAAGACCGTG TTATAGTACA ACAGCATCAG GTTGTTAGTG   532
CHV       TCATATAAAA CAAGAACGTG TTATAGTACA ACAGCATCAG GTTGTTAGTG   541
           ^^ ^ ^ ^^ ^^^
```

```
LEPP      CTAGAACACA AAATTATTAC CCAGAGTTCA GCATCGCTGT AC-TTTTGTA   587
AR310     CTAGAACACA AAATTATTAC CCAGAGTTCA GCATCGCTGT AC-TTTTGTA   587
ISU-1     ---------- ---------- ---------- ---------- ----------   329
1894      CTAGAGCACA AAATTATTAT CCAGAGTTCA GCATCGCTGT ACTTTTTGTA   560
725       CTAGAACACA AAATTATTAC CCAGAGTTCA GCATCGCTGT AC-TTTTGTA   587
306       CTAGAACACA AAACTATTAC CCAGAGTTCA GCATCGCTGT ACTCTTTGTA   590
5170      CTAGAACACA AAATTATTAC CCAGAGTTCA GCATCGCTGT ACTTTTTGTA   588
PON       CTAGAACACA AAATTATTAC CCAGAGTTCA GCATCGCTGT ACTTTTTGTA   582
CHV       CTAGAACACA AAATTATTAC CCAGAGTTCA GCATCGCTGT ACTTTTTGTA   591
```

Figure 4.(continued)

```
LEPP      TCTTTCCTAG CTTTGTACCG TAGTACAAAC TTTAAGACGT GTGTCGGTAT   637
AR310     TCTTTCCTAG CTTTGTACCG TAGTACAAAC TTTAAGACGT GTGTCGGTAT   637
ISU-1     ---------- ---------- ---------- -TTAAGACGT GTGTCGGTAT   348
1894      TCTTTTCTAG CTTTGTACCG CAGTCCAAAC TTTAAGACGT GTGTCGGTAT   610
725       TCTTTCCTAG CTTTGTACCG TAGTACAAAC TTTAAGACGT GTGTCGGTAT   637
306       TCTTTTCTAG CTTTGTACCG TAGTACAAAC TTTAAGACGT GTGTCGGTAT   640
5170      TCTTTCTTAG CTTTTTACCG TAGTACAAAC TTTAAGACGT GTGTCGGTAT   638
PON       TCTTTCCTAG CTTTGTACCG TAGTACAAAC TT--AGACGT GTGTCGGTAT   630
CHV       TCTTTTCTAG CTTTGTACCG TAGTACAAAC TTTAAGACGT GTGTCGGCAT   641
                                          ^^^^^^ ^^^^^^^ ^^
```

```
                                    Primer  048
                                  ========== ========== =====
LEPP      CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   687
AR310     CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   687
ISU-1     CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   398
1894      CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   660
725       CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   687
306       CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   690
5170      CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   688
PON       CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   680
CHV       CTTAATGTTT AAGATTTTAT CAATGACACT TTTAGGACCT ATGCTTATAG   691
          ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^
```

Figure 4.(continued)

```
LEPP      TATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTGT CTTATCTTTA   737
AR310     TATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTGT CTTATCTTTA   737
ISU-1     TATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTTT CTTATCTTTA   448
1894      TATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTGT CTTATCTTTA   710
725       TATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTGT CTTATCTTTA   737
306       CATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTGT CTTATCTTTA   740
5170      TATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTGT CTTATCTTTA   738
PON       TATATGGTTA CTACATTGAT GGCATTATTA CAACAACTGT CTTATCTTTA   730
CHV       CATATGGTTA CTACATTGAT GGCATTGTTA CAACAACTGT CTTATCTTTA   741
            ^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^ ^ ^^^^^^^^^^


LEPP      AGATTCGCCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAATT   787
AR310     AGATTCGCCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAATT   787
ISU-1     AGATTCGCCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAATT   498
1894      AGATTCGCCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAATT   760
725       AGATTCGCCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAATT   787
306       AGATTTGTCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAATT   790
5170      AGATTCGCCT ACTTAGCATA CTTTTGGTAT GTTAACAGTA GGTTTGAATT   788
PON       AGATTTGCCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAATT   780
CHV       AGATTTGCCT ACTTAGCATA CTTTTGGTAT GTTAATAGTA GGTTTGAAGT   791
          ^^^^^ ^ ^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^ ^^^^ ^^^^^^^^ ^
```

Figure 4.(continued)

```
LEPP      TATTTTATAC AACACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   837
AR310     TATTTTATAC AACACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   837
ISU-1     TATTTTATAC AACACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   548
1894      TATTTTATAC AACACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   810
725       TATTTTATAC AACACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   837
306       TATTTTATAC AATACAACGA CACTCATGTT TGTACATGGC AGAGCTGTAC   840
5170      TATTTTATAC AACACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   838
PON       TATTTTATAC AACACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   830
CHV       TATTTTATAC AATACAACGA CACTCATGTT TGTACATGGC AGAGCTGCAC   841
          ^^^^^^^^^^ ^^ ^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^ ^^

LEPP      CGTTTAAGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT ATATGGTGGC   887
AR310     CGTTTAAGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT ATATGGTGGC   887
ISU-1     CGTTTAAGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT ATATGGTGGC   598
1894      CGTTTAAGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT GTATGGTGGC   860
725       CGTTTAAGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT ATATGGTGGC   887
306       CGTTTATGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT GTATGGTGGC   890
5170      CGTTTAAGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT ATATGGTGGC   888
PON       CGTTTAAGAG AAGTTCTCAC AGCTTTATTT ATGTCACATT ATATGGTGGC   880
CHV       CGTTTAAGAG AAGTTCTCAC AGCTCTATTT ATGTCACATT GTATGGTGGC   891
          ^^^^^^ ^^^ ^^^^^^^^^^ ^^^^ ^^^^^ ^^^^^^^^^^  ^^^^^^^^^
```

Figure 4.(continued)

```
LEPP      ATAAATTATA TGTTTGTGAA TGACCTCACG TTGCATTTTG TAGACCCTAT   937
AR310     ATAAATTATA TGTTTGTGAA TGACCTCACG TTGCATTTTG TAGACCCTAT   937
ISU-1     ATAAATTATA TGTTTGTGAA TGACTTCATG TTGCATTTTG TAGACCCTAT   648
1894      ATAAATTATA TGTTTGTGAA TGACCTCACG TTGCATTTTG TAGACCCTAT   910
725       ATAAATTATA TGTTTGTGAA TGACCTCACG TTGCATTTTG TAGACCCTAT   937
306       ATAAATTATA TGTTTGTGAA TGACCTCACG TTGCATTTTG TAAACCCTAT   940
5170      ATAAATTATA TGTTTGTGAA TGACCTCACG TTGCATTTTG TAGACCCTAT   938
PON       ATAAATTATA TGTTTGTGAA TGACTTCACG TTGCATTTTG TAGACCCTAT   930
CHV       ATAAATTATA TGTTTGTGAA TGACCTCACG TTGCATTTTG TAGACCCTAT   941
          ^^^^^^^^^^ ^^^^^^^^^^ ^^^^ ^^^ ^ ^^^^^^^^^^ ^^ ^^^^^^^


LEPP      GCTTGTAAGC ATAGCAATAC GTGGCTTAAC TCATGCTGAT CTAACTGTAG   987
AR310     GCTTGTAAGC ATAGCAATAC GTGGCTTAAC TCATGCTGAT CTAACTGTAG   987
ISU-1     GCTTGTAAGC ATAGCAATAC GTGGCTTAAC TCATGCTGAT CTAACTGTAG   698
1894      GCTTGTAAGC ATAGCAATAC GTGGCTTAGC TCATGCTGAT CTAACTGTAG   960
725       GCTTGTAAGC ATAGCAATAC GTGGCTTAAC TCATGCTGAT CTAACTGTAG   987
306       GCTTGTAAGC ATAGCAATAC GTGGCTTAGC TCATGCTGAT CTAACTGTAG   990
5170      GCTTGTAAGC ATAGCAATAC GTGGCTTAAA TCATGCTGAT CTAACTGTAG   988
PON       GCTTGTAAGC ATAGCAATAC GTGGCTTAAC TCATGCTGAT CTAACTGTAG   980
CHV       GCTTGTAAGC ATAGCAATAC GTGGCTTAGC TCATGCTGAT CTAACTGTAG   991
          ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^   ^^^^^^^^^^ ^^^^^^^^^^
```

Figure 4.(continued)

```
LEPP      TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATAT ATTTTCACAG   1037
AR310     TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATAT ATTTTCACAG   1037
ISU-1     TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATAT ATTTTCACAG   748
1894      TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATAT ATTTTCACAG   1010
725       TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATAT ATTTTCACAG   1037
306       TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATGT ATTTTCACAG   1040
5170      TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATAT ATTTTCACAG   1038
PON       TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATAT ATTTTCACAG   1030
CHV       TTAGAGCAGT TGAACTTCTC AATGGTGATT TTATTTATGT ATTTTCACAG   1041
          ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^ ^ ^^^^^^^^^^

                                                      *****
LEPP      GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   1087
AR310     GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   1087
ISU-1     GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   798
1894      GATTCTGTAG TTGGTGTTTA CAATGCAGCC TTTTCTCAGG CAGTTCTAAA   1060
725       GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   1087
306       GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   1090
5170      GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   1088
PON       GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   1080
CHV       GAGCCCGTAG TCGGTGTTTA CAATGCAGCC TTTTCTCAGG CGGTTCTAAA   1091
          ^^    ^^^^ ^ ^^^^^^^^ ^^^^^^^^^^ ^^^^^^^^^^ ^ ^^^^^^^^
```

Figure 4.(continued)

```
                                                    ORF  4
          **                                        +1→
LEPP      CGAAATTGAC TTAAAAGAAG AAGAGGGAGA CCGTACCTAT GACGTTTCCT   1137
AR310     CGAAATTGAC TTAAAAGAAG AAGAGGGAGA CCGTACCTAT GACGTTTCCT   1137
ISU-1     CGAAATTGAC TTAAAAGAAG AAGAGGGAGA CCGTACCTAT GACGTTTCCT   848
1894      CGAAATTGAC TTAAAAGAAG AAGAGGGAGA CCGTACCTAT GACGTTTCCT   1110
725       CGAAATTGAC TTAAAAGAAG AAGAGGGAGA CCGTACCTAT GACGTTTCCT   1137
306       CGAAATTGAC TTAAAAGAAG AAGAAGAAGA CCATTCCTAT GACGTTTCCT   1140
5170      CGAAATTGAC TTAAAAGAAG AAAAGGGAGA CCGTACCTAT GACGTTTCCT   1138
PON       CGAAATTGAC TTAAAAGAAG AAGAGGGAGA CCGTACCTAT GACGTTTCCT   1130
CHV       CGAAATTGAC TTAAAAGAAG AAGAAGAAGA CCGTACCTAT GACGTTTCCT   1141
          ^^^^^^^^^^ ^^^^^^^^^^ ^^ ^ ^ ^^^ ^^ ^ ^^^^^ ^^^^^^^^^^

          ORF  3-1                          Primer  622
          Stop                           =====  ========
LEPP      AGGGCATTGA CTGTCATAGA TGACAATGGA ATGGTCATT     1185
AR310     AGGGCATTGA CTGTCATAGA TGACAATGGA ATGGTCATT     1185
ISU-1     AGGGCATTGA CTGTCATAGA TGATAATGGA ATGGTCATT     896
1894      AGGGCATTGA CTGTCATAGA TGACAATGGA ATGGTCATT     1158
725       AGGGCATTGA CTGTCATAGA TGACAATGGA ATGGTCATT     1176
306       AGGGCATTGA CTGTCATAGA TGACAATGGA ATGGTCATT     1179
5170      AGGGCATTGA CTGTCATAGA CGACAATGGA ATGGTCATT     1177
PON       AGGGCATTGA CTGTCATAGA TGATAATGGA ATGGTCATT     1169
CHV       AGGGCATTGA CTGTCATAGA TGACAATGGA ATGGTCATT     1189
          ^^^^^^^^^^ ^^^^^^^^^^  ^^ ^^^^^^ ^^^^^^^^^
```

Figure 4.(continued)

CONCLUSION

In our study we characterized four new isolates of PRCV, 725, 306, PON, and 5170 NVSL. All of these isolates were procured from respiratory extracts except for isolate 5170 NVSL, which was isolated from a fecal sample along with TGEV. PCR amplification of the 5' end of the S gene of these isolates yielded an approximately 200 base pair product which is indicative of PRCV. Upon sequencing, large deletions in this region characteristic of PRCV were seen. These four new isolates were shown to vary in the size of the S gene deletion. These new isolates will be helpful in understanding coronavirus pathogenicity mechanisms.

In another portion of our experiment the nucleotide sequences of genes S, 3 and 3-1 were determined and compared to other PRCV isolates and TGEV isolate CHV. All of the isolates had a large, inframe deletion in the 5' end of the S gene. PRCV isolate 725 had an identical deletion to isolates AR310 and LEPP of 621 nucleotides which started 47 nucleotides after the S gene start site. According to previous studies, these are the smallest deletions that have been recorded in this region. The deletion in other US and European isolates characterised to date range from 672 to 681 nucleotides in length. PON also had a deletion which was the same as a previously characterized isolate, ISU-1. They both had a deletion which is 681 nucleotides in length and begins 62 nucleotides after the start site. Isolates 306 and 5170 NVSL both had deletions which are specific to those isolates. Isolate 306 has a 675 nucleotide deletion in the 5' end of the S gene which starts 57 nucleotides after the S gene start site.

Isolate 5170 NVSL is very unique as it has a 711 nucleotide deletion which begins only 27 nucleotides after the S gene start site. This is the largest deletion which has been characterized so far and also starts the closest to the S gene start site. It is possible, due to the location of this deletion that the virus may be missing a good portion of the signal peptide. This may be important in the tissue tropsim of this isolate. The herd that 5170 NVSL was isolated from showed chronic diarrhea and the virus itself was isolated from a fecal sample. TGEV was also isolated from this sample. Further studies should be done to determine the relevance of this deletion in tissue tropism and pathogenicity.

Sequence analysis of the gene 3 of the four PRCV isolates showed some commonality with previously studied isolates, but also some unique diversity. The gene 3 of all of the new isolates contained the CTAAAC leader binding site. The isolates 725 and 5170 NVSL were predicted to yield a protein of 72 amino acids which is the same as PRCV isolates AR310 and LEPP as well as the virulent Miller strain of TGEV. The PRCV isolate PON has the CTAAAC leader binding site, but is predicted to yield a truncated protein 3 of only 17 amino acids due to a 1 nucleotide change which codes for a stop site. Isolate 306 has several changes from the others in this region. It has a deletion of 7 nucleotides in this region which causes a frameshift and it also has a 28 nucleotide insertion which codes for a stop site. The predicted protein from this region is expected to be only 67 amino acids.

Analysis of the 3-1 genes of the four PRCV isolates showed that there was a large amount of variation in this region among the isolates. The gene 3-1 of the PRCV isolates 725, PON, and 5170 NVSL were all preceded by the CTAAAC leader RNA binding site. Isolate 306 had no leader binding site and no

3-1 start site due to a 52 nucleotide deletion in this region. It is predicted that protein 3-1 is not made by this isolate. This is a unique feature of this isolate. It has been theorized that the gene 3 codes for a non-structural protein which is important in pathogenesis. The variations seen among these isolates in this region may yield answers to the mechanism of pathogenesis.

In looking at the biological characteristics of these isolates it is possible to see that they have several factors in common. They all replicate in the ST cell line, as well as the IPEC-1 cell line. It is interesting to note that all isolates except 1894 caused a cytopathic infection. The unique feature of this virus is its small deletion in the S gene and also its 23 nucleotide deletion in the 3 gene. It is possible that this small deletion can influence the pathogenicity of this virus, as previous studies have shown to not be very virulent. This isolate also showed the smallest average plaque size.

Additional studies should be done to clarify the role that the genes 3 and 3-1 play in the virulence of PRCV and TGEV. With the new isolates that we have characterized, we now have constructs with several varying deletions and additions in these regions of their genomes. Animal studies could be done to see if there is any correlation between the changes in these regions and the pathogenicity of these viruses.

BIBLIOGRAPHY

Baric, R. S., K. Fu, W. Chen, and B. Yount. 1995. High recombination and mutation rates in mouse hepatitis virus suggest that coronaviruses may be potentially important emerging viruses. Pages 571-576 in P.J. Talbot and G.A. Levy, ed. Corona- and Related Viruses. Plenum Press, New York, NY.

Bernard, S., E. Bottreau, J. M. Aynaud, P. Have and L. Szymansky. 1989. Natural infection with the porcine respiratory coronvavirus induces protective lactogenic immunity against transmissible gastroenteritis virus. Vet. Microbiol. 21:1-8.

Bohl, E. H. 1989. Transmissible gastroenteritis virus (classical enteric variant). Pages 139-153 in M. B. Pensaert, ed. Virus infections of porcines. Elsevier Science Publishing Company Inc., New York, NY.

Britton, P., K. L. Mawditt and K. W. Page. 1991. The cloning and sequencing of the virion protein genes from a British isolate of porcine respiratory coronavirus: comparison with transmissible gastroenteritis virus genes. Virus Research. 21:181-198.

Callebaut, P., I. Correa, M. Pensaert, G. Jimenez and L. Enjuanes. 1988. Antigenic differentiation between transmissible gastroenteritis virus of swine and a related porcine respiratory coronavirus. Journal of Virology. 69:1725-1730.

Callebaut, P., M. B. Pensaert, and J. Hooyberghs. 1989. A competitive inhibition ELISA for the differentiation of serum antibodies from pigs infected with transmissible gastroenteritis virus (TGEV) or with the TGEV-related porcine respiratory coronavirus. Veterinary Microbiology. 20:9-19.

Cheever, F. S., J. B. Daniels, A. M. Pappenheimer, and O. T. Bailey. 1949. A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. Isolation and biologic properties of the virus. Journal of Experimental Medicine. 90:181-194.

Cox, E., L. Hooyberghs and M. B. Pensaert. 1990a. Sites of replication of a porcine respiratory coronavirus related to transmissible gastroenteritis virus. Research in Veterinary Science. 48:165-169.

Cox, E., M. B. Pensaert, P. Callebaut and K. Van Deun. 1990b. Intestinal replication of a porcine respiratory coronavirus closely related antigenically to the enteric transmissible gastroenteritis virus. Veterinary Medicine. 23:237-243.

Cox, E., M. B. Pensaert and P. Callebaut. 1993. Intestinal protection against challenge with transmissible gastroenteritis virus of pigs immune after infection with porcine respiratory coronavirus. Vaccine. 11:267-272.

Delmas, B., J. Gelfi, and R. L'Haridon. 1990. Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. Nature. 357:417-420.

Elazhary, Dr. Y., and others. 1992. Porcine respiratory coronavirus isolated from young piglets in Quebec. The Veterinary Record. May 30, 1992 p.500.

Garwes, D. J., D. H. Pocock, and B. V. Pike. 1976. Isolation of subviral components from transmissible gastroenteritis virus. Journal of General Virology. 32:283-294.

Garwes, D. J., F. Stewart, S. F. Cartwright, I. Brown. 1988. Differentiation of porcine coronavirus from transmissible gastroenteritis virus. The Veterinary Record. January 23, 1988, p.86-87.

Garwes, D. J., M. H. Lucas, D. A. Higgins, B. V. Pike, and S. F. Cartwright. 1978. Antigenicity of structural components from porcine transmissible gastroenteritis virus. Veterinary Microbiology. 3:179-190.

Godet, M., R. L'Haridon, J. F. Vautherott, and H. Laude. 1992. TGEV coronavirus ORF 4 encodes a membrane protein that is incorporated into virions. Virology. 188:666-675.

Godet, M., D. Rasschaert, and H. Laude. 1991. Processing and antigenicity of entire and anchor-free spike glycoprotein S of coronavirus TGEV expressed by recombinant baculovirus. Virology. 185:732-740.

Halbur, P., P. S. Paul, M. Frey, I. Morozov, and T. Sirinarmitr. 1995. TGE-like disease outbreak: A foreign disease or a variant?. 1995 Allen D. Leman Swine Conference, p.1-2,

Halbur, P., P. S. Paul, E. M. Vaughn, and J. J. Andrews. 1993. Experimental reproduction of pneumonia in gnotobiotic pigs with porcine respiratory coronavirus AR310. Journal of Veterinary Diagnostic Investigation. 5:184-188.

Hill, H., J. Biwer, R. Woods, and R. Wesley. 1989. Porcine respiratory coronavirus isolated from two U.S. swine herds. Pages 333-335 in T. A. Neuzil, ed. Proceedings of the American Association of Swine Practitioners. American Association of Swine Practitioners, Des Moines, IA.

Holmes, K. V. 1990. Coronaviridae and their replication. Pages 841-856 in B. N. Fields, D. M. Knipe, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman, eds. Fields virology. 2nd edition. Raven Press, New York, NY.

Ingersoll J. D., and D. E. Wylie. 1988. Identification of viral antigens that induce antibody responses on exposure to coronaviruses. American Journal of Veterinary Research. 49:1467-1471.

Jabrane, A., C. Girard, and Y. Elazhary. 1994. Pathogenicity of porcine respiratory coronavirus isolated in Quebec. Canadian Veterinary Journal. 35:86-92.

Jabrane, A., Y. Elazhary, B. G. Talbot, R. Ethier, C. Dubuc, R. Assaf. 1992. Porcine respiratory coronavirus in Quebec:Serological studies using a competitive inhibition enzyme-linked immunosorbent assay. Canadian Veterinary Journal. 33:727-733.

Kapke, P. A., and D. Brian. 1986. Sequence analysis of the porcine transmissible gastroenteritis coronavirus nucleocapsid protein gene. Virology. 151:41-49.

Kapke, P. A., F. Y. T. Tung, B. G. Hogue, D. A. Brian, R. D. Woods, and R. Wesley. 1988. The amino-terminal signal peptide on the porcine transmissible gastroenteritis coronavirus matrix protein is not an absolute requirement for membrane translocation and glycosylation. Virology. 165:367-376.

Lai, M. M. C. 1990. Coronavirus: organization, replication, and genome expression of genome. Annual Reviews in Microbiology. 44:303-333.

Lanza, I., I. H. Brown, and D. J. Patton. 1992. Pathogenicity of concurrent infection of pigs with porcine respiratory coronavirus and swine influenza virus. Research in Veterinary Science. 53:309-314.

Laude, H., J. Gelfi, L. Lavenant, and B. Charley. 1992. Single amino acid changes in the viral glycoprotein M affect induction of alpha interferon by the coronavirus transmissible gastroenteritis virus. Journal of Virology. 66:743-749

Laude, H., K. Van Reeth, and M. Pensaert. 1993. Porcine respiratory coronavirus: molecular features and virus-host interactions. Veterinary Research. 24:125-150.

McClurkin, A. W., and J. O. Norman. 1966. Studies of transmissible gastroenteritis of swine. II. Selected characterisitics of a cytopathogenic virus common to five isolates from transmissible gastroenteritis. Canadian Journal of Comparative Medicine. 34:347-349.

McIntosh, K. 1990. Coronaviruses. Pages 857-864 in B. N. Fields, D. M. Knipe, R. M. Channock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman, eds. Fields virology. 2nd edition. Raven Press, New York, NY.

McIntosh, K., W. B. Becker, and R. M. Chanock. 1967. Growth in suckling mouse brain of "IBV-like" viruses from patients with upper respiratory tract disease. Proceedings of the National Academy of Science. 58:2268-2273.

Mendez, A., C. Smerdou, A. Izeta, F. Gebauer, and L. Enjuanes. Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. Virology. 217:495-507.

Moon, H. W. 1978. Mechanisms in the pathogenesis of diarrhea: A review. Journal of the American Veterinary Medicine Association. 172:443-448.

O'Toole, D., I. Brown, A. Bridges, and S. F, Cartwright. 1989. Pathogenicity of experimental infection with 'pneumotropic' porcine coronavirus. Research in Veterinary Science. 47:23-29.

Page, K. W., K. L. Mawditt, and P. Britton. 1991. Sequence comparison of the 5' end of the mRNA 3 from transmissible gastroenteritis virus and porcine respiratory coronavirus. Journal of General Virology. 72:570-587.

Paul, P. S., E. M. Vaughn, and P. G. Halbur. 1992. Characterization and pathogenicity of a new porcine respiratory coronavirus strain AR310. Proceedings of the International Pig Veterinary Society Congress. 12:92.

Pensaert, M. B. 1989. Transmissible gastroenteritis virus (respiratory varient). Pages 154-165 in M. B. Pensaert, ed. Virus infections of porcines. Elsevier Science Publishing Company Inc., New York, NY.

Pensaert, M. B., P. DeBouck, and D. J. Reynolds. 1981: An immunoelectron microscopic and immunofluorescent study on the antigenic relationship between the coronavirus-like agent, CV 777, and several coronaviruses. Archives of Virology. 68:45-52.

Pensaert, M. B., P. Callebaut, and J. Vergote. 1986. Isolation of a porcine respiratory, non-enteric coronavirus related to transmissible gastroenteritis virus. Veterinary Quarterly. 8:257-261.

Pensaert, M. B., E. Cox, K. Van Deun, and P. Callebaut. 1993. A sero-epizootological study of porcine respiratory coronavirus in Belgian swine. Veterinary Quarterly. 15:16-20.

Rasschaert, D., M. Duarte, and H. Laude. 1990. Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. Journal of General Virology. 71:2599-2607.

Risco, C., I. M. Anton, C. Sune, A. M. Pedregosa, J. M. Martin-Alonso, F. Parra, J. L. Carrascosa, and L. Enjuanes. 1995. Membrane protein molecules of transmissible gastroenteritis coronavirus also expose the carboxyl-terminal region on the external surface of the virion. Journal of Virology. 69:5269-5277.

Rossen, J. W., C. P. J. Bekker, W. F. Voorhout, G. J. A. M. Strous, A. Van Der Ende, and P. J. M. Rottier. 1994. Entry and release of transmissible gastroenteritis coronavirus are restricted to apical surfaces of polarized epithelial cells. Journal of Virology. 68:7966-7973.

Saif, L., and E. H. Bohl. 1986. Transmissible gastroenteritis. Pages 255-274 in A. D. Leman, R. D. Glock, W. L. Mengeling, R. H. C. Penny, E. Scholl, and B. Straw, ed. Diseases of swine, 6th edition. Iowa State Universtiy Press, Ames, IA.

Saif, L., and R. D. Wesley. 1992. Transmissible gastroenteritis. Pages 362-386 in A. D. Leman, B. E. Strauss, W. L. mengling, S. D'Allaire, D. J.

Taylor ed. Diseases of Swine, 7th edition. Iowa State University Press, Ames, IA.

Sanchez, C. M., F. Gebauer, C. Sune, A. Mendez, J. Dopazo and L. Enjuanes. 1992. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. Virology. 190:92-105.

Sanchez, C. M., G. Jimenez, M. D. Laviada, I. Correa, C. Sune, M J. Bullido, F. Gerbauer, C. Smerdou, P. Callebaut, J. M. Escribano, and L. Enjuanes. Antigenic homolgy among coronaviruses related to transmissible gastroenteritis virus. Virology. 174:410-417.

Simkins, R. A., P. A. Weilnau, J. Bias, L. Saif. 1992. Antigenic variation among transmissible gastroenteritis virus (TGEV) and porcine respiratory coronavirus strains detected with monoclonal antibodies to the S protein of TEGV. American Journal of Veterinary Research. 53:1253-1258.

Simkins, R. A., P. A. Weilnau, J. Van Cott, T. A. Brims, and L. Saif. 1993. Competition ELISA, using monoclonal antibodies to the transmissible gastroenteritis virus (TGEV) S protein, for serologic differentiation of pigs infected with TGEV or porcine respiratory coronavirus. American Journal of Veterinary Research. 54:254-259.

Sirinarumitr, T., P. S. Paul, J. P. Kluge, and P. G. Halbur. 1996. In situ hybridization for the detection of swine enteric and respiratory coronaviruses, transmissible gastroenteritis virus (TGEV) and porcine respiratory coronavirus (PRCV), in formalin-fixed paraffin-embedded tissues. Journal of Virological Methods. 56:149-160.

Spaan, W., D. Cavanagh, and M. C. Horzinek. 1988. Coronaviruses: Structure and genome expression. Journal of General Virology. 69:2939-2952.

Tung, F. T., S. Abraham, M. Sethna, S. L. Hung, P. Sethna, B. G. Hogue, and D. Brian. 1992. The 9-kDa protein encoded at the 3' end of the porcine transmissible gastroenteritis coronavirus genome is membrane-associated. Virology. 186:676-683.

Tyrrell, D. A. J., and M. L. Bynoe. 1965. Cultivation of a novel type of common cold virus in organ cultures. British Medical Journal. 1:1467-1470.

Tyrrell, D. A. J., J. D. Almeida, and D. M. Berry. 1968. Coronaviruses. Nature. 220:650.

Van Marle, G., W. Luytjes, R. G. Van Der Most, T. Van Der Straaten, and W. J. M. Spaan. 1995. Regulation of coronavirus mRNA transcription. Journal of Virology. 69:7851-7856.

Van Nieuwstadt, A. P., and J. M. A. Pol. 1989. Isolation of a TGE virus-related respiratory coronavirus causing fatal pneumonia in pigs. Veterinary Record. 124:43-44.

Van Nieuwstadt, A. P., and J. Boonstra. 1992. Comparison of the antibody response to transmissible gastroenteritis virus and porcine respiratory coronavirus, using monoclonal antibodies to antigenic sites A and X of the S glycoprotein. American Journal of Veterinary Research. 53:184-190.

Van Reeth, K., and M. B. Pensaert. 1994. Porcine respiratory coronavirus-mediated interference against influenza virus replication in the respiratory tract of feeder pigs. American Journal of Veterinary Research. 55:1275-1281.

Vaughn, E. M., P. G. Halbur, and P. S. Paul. 1995. Sequence comparison of porcine respiratory coronavirus isolates reveals heterogeneity in the S, 3, and 3-1 genes. Journal of Virology. 69:3176-3184.

Vaughn, E. M., P. G. Halbur, and P. S. Paul. 1994. Three new isolates of porcine respiratory coronavirus with various pathogenicities and spike (S) gene deletions. Journal of Clinical Microbiology. 32:1809-1812.

Vaughn, E. M., and P. S. Paul. 1993. Antigenic and biological diversity among transmissible gastroenteritis virus isolates of swine. Veterinary Microbiology. 36:333-347.

Weingartl, H., and J. B. Derbyshire. 1993. Binding of porcine transmissible gastroenteritis virus by enterocytes from newborn and weaned piglets. Veterinary Microbiology. 35:223-232.

Wesley, R. D., R. D. Woods, and A. K. Cheung. 1990a. Genetic basis for the pathogenesis of transmissible gastroenteritis virus. Journal of Virology. 64:4761-4766.

Wesley, R. D., I. V. Wesley, and R. D. Woods. 1991. Differentiation between transmissible gastroenteritis virus and porcine respiratory coronavirus using cDNA probe. Journal of Veterinary Diagnostic Investigation. 3:29-32.

Wesley, R. D., R. D. Woods, H. H. Hill, and J. D. Biwer. 1990b. Evidence for a porcine respiratory coronavirus antigenically similar to transmissible gastroenteritis virus, in the United States. Journal of Veterinary Diagnositic Investigation. 2:312-317.

Wesley, R. D., R. D. Woods, and A. K. Cheung. 1991. Genetic analysis of porcine respiratory coronavirus, an attenuated variant of transmissible gastroenteritis virus. Journal of Virology. 65:3369-3373.

Woods, R. D. 1978. Small plaque variant transmissible gastroenteritis virus. Journal of the American Veterinary Medicine Association. 173:643-647.

Woods, R. D., N. F. Cheville, and J. E. Gallagher. 1981. Lesions in the small intestine of newborn pigs inoculated with porcine, feline, and canine coronaviruses. American Journal of Veterinary Research. 42:1163-1169.

Zhu, X. L., P. S. Paul, E. Vaughn, and A. Morales. 1990. Characterization and reactivity of monoclonal antibodies to the Miller strain of transmissible gastroenteritis virus of swine. American Journal of Veterinary Research. 51:232-238.